

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 03:30:48 ; Search time 3690 Seconds  
(without alignments)  
2981.185 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111950\_8301/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bite -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10042894 @CGN\_1\_1\_3437\_@runat\_15062005\_111950\_8301 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1          | 1521   | 100.0       | 960    | 8  | CC344833    |
| 2          | 1406.5 | 92.5        | 923    | 9  | CG240397    |
| 3          | 1256   | 82.6        | 818    | 9  | CG453578    |
| 4          | 1248   | 82.1        | 873    | 8  | CC344824    |
| 5          | 1183   | 77.8        | 867    | 9  | CC724960    |
| 6          | 1149   | 75.5        | 719    | 9  | CG284572    |
| 7          | 1123   | 73.8        | 801    | 6  | CA134480    |
| 8          | 1034   | 68.0        | 656    | 6  | CA245555    |
| 9          | 993.5  | 65.3        | 722    | 6  | CA130685    |

|    |       |      |      |   |          |
|----|-------|------|------|---|----------|
| 10 | 992   | 65.2 | 591  | 9 | CG776236 |
| 11 | 987   | 64.9 | 671  | 6 | CA202412 |
| 12 | 981   | 64.5 | 3374 | 3 | AY109355 |
| 13 | 976   | 64.2 | 722  | 9 | CC724950 |
| 14 | 932   | 61.3 | 744  | 6 | CA130686 |
| 15 | 926   | 60.9 | 1191 | 7 | CK211413 |
| 16 | 921.5 | 60.6 | 593  | 8 | BZ774817 |
| 17 | 920   | 60.5 | 960  | 9 | CG073495 |
| 18 | 903   | 59.4 | 955  | 8 | CC384766 |
| 19 | 869.5 | 57.2 | 794  | 2 | BF617713 |
| 20 | 811.5 | 53.4 | 907  | 9 | CG240387 |
| 21 | 807   | 53.1 | 575  | 9 | CG734025 |
| 22 | 770   | 50.6 | 619  | 6 | CA204213 |
| 23 | 758   | 49.8 | 1099 | 7 | CK215908 |
| 24 | 727   | 47.8 | 1123 | 7 | CK214494 |
| 25 | 724.5 | 47.6 | 588  | 5 | BU999708 |
| 26 | 724   | 47.6 | 910  | 7 | CK253394 |
| 27 | 714   | 46.9 | 551  | 8 | BH408067 |
| 28 | 705.5 | 46.4 | 1119 | 3 | CNSOALFK |
| 29 | 666.5 | 43.8 | 1007 | 3 | CNSO9Z6Q |
| 30 | 664.5 | 43.7 | 1813 | 3 | CNSOALNR |
| 31 | 644   | 42.3 | 557  | 2 | BF277833 |
| 32 | 642   | 42.2 | 831  | 9 | CG553151 |
| 33 | 640.5 | 42.1 | 1431 | 3 | CNSO9ZLQ |
| 34 | 640   | 42.1 | 620  | 5 | BQ620191 |
| 35 | 617   | 40.6 | 804  | 7 | CNS23418 |
| 36 | 613   | 40.3 | 453  | 4 | B1956405 |
| 37 | 606   | 39.8 | 433  | 4 | BM099879 |
| 38 | 591   | 38.9 | 883  | 7 | CK934794 |
| 39 | 585.5 | 38.5 | 752  | 7 | CO865620 |
| 40 | 584   | 38.4 | 448  | 8 | CC028437 |
| 41 | 581.5 | 38.2 | 478  | 4 | BG417083 |
| 42 | 580   | 38.1 | 876  | 9 | CG073494 |
| 43 | 567   | 37.3 | 409  | 4 | BM099878 |
| 44 | 562   | 36.9 | 927  | 9 | CG284580 |
| 45 | 559   | 36.8 | 431  | 9 | CG776505 |

#### ALIGNMENTS

RESULT 1  
CC344833

LOCUS

CGIAQ20TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0359C15,  
genomic survey sequence.

ACCESSION CC344833

VERSION CC344833.1

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE

AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

COMMENT

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..960

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

FEATURES  
source

```
/clone="ZMMBMA0359C15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 2,95e-150 Length: 960
Score: 1521.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-042-894A-8 (1-289) x CC344833 (1-960)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys 20
Db 18 ATGTCGACCTCACCCGCGGAGCACCAAGTCGCCGCCACCGCCCTCCGCGCAG 77
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArg 40
Db 78 CTGGGCCGCTCATCGACGGCTCCGCCCTCTTCTACAAGCCGCTCCAGGCCGCGCACCGT 137
QY 41 GlyGluHisGluValAlaPheTyrlsGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 138 GGGAGACGAGGTGCGCTTCTATAGGGCTTCTCCGCCACCGCCCTCCGCGCCGC 197
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 198 ATCCGAGACACCTTCTTCCCGGTTCCACGCGACGACTCTCTCCACCGCGCGCAG 257
QY 81 ProGlyGluProHisProHisValLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 258 CCGGGGAGCGCATCTCACCTCGCTTCGACGACTCTCTCGCGGGTTTCAGGCGCC 317
QY 101 CysValAlaAspIleIysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 318 TCGGTCCGACATCAAGATCGCGCCATCATCGTGGCCACCGAGTTTCGCGGAGCCCTAC 377
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 378 ATCCCAAGTGCTCGCCCAAGGACCGCGGACCCACGAGCGTCTCGATTCGCGTC 437
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 438 TCCGGGTCGAGTCTCGGCCCGGAGGCGCGTGTGGCGGACGAGCGCGCGAGGTG 497
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
Db 498 AAGGCCATGGACACCCCGCGGTGCGCGCGGTCTCCGCGGTACGTGTCATCCGTTGCC 557
QY 181 AspGluGlyMetAspCysAlaLeuAlaValTrpGlyGlyLysGlyGlyValLeu 200
Db 558 GACGAGGGATGAGCTGTGGCTCGCGCGCGGTGTACGGAGAAAAGTGGAGTCTTG 617
QY 201 SerGlnLeuArgGluLeuIysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220
Db 618 TCACAGCTCGCGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCACACTTCTACTCG 677
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyValGlyGly 240
Db 678 GGTCTGATTTCTTGGGCTATGATGTCTGTCAGTCCGAGCGCGGAGATGGGGGTGG 737
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
Db 738 GTGACGGTGAAGCTGTGTGACTTTGCCATGTGGCCGAGGGTGTATGGGCTGATTCAC 797
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
Db 798 AACTTCTCGGGGGGCTCTGCTCGTGTATCAAGTTCTGTTTCTGACATTTCTCGGAGACT 857
QY 281 ProHisThrGlnProLeuGlyProSer 289
```

```
|||||
858 CCTCATACGACGCTTTGGTCTTCT 884

RESULT 2
CG240397
LOCUS
DEFINITION
CG240397 ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0788E05,
genomic survey sequence.
ACCESSION
CG240397
VERSION
CG240397.1 GI:34140283
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 923)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reinick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGYCI27TH
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..923
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0788E05"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Alignment Scores:
Pred. No.: 3,73e-138 Length: 923
Score: 1406.50 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 92.47% Indels: 2
DB: 9 Gaps: 0

US-10-042-894A-8 (1-289) x CG240397 (1-923)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys 20
Db 34 ATCCCGACCTTCCACCCCGGAGCACCAAGTCGCGGTCCCGCGCTCCCGCAGCAG 93
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArg 40
Db 94 CTGGGCCGCTCATCGACGGCTCCGCCCTCTTCTACAAGCCGCTCCAGGCCGCGCACCGT 153
QY 41 GlyGluHisGluValAlaPheTyrlsGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 154 GGGAGACGAGGTGCGCTTCTATAGGGCTTCTCCGCCACCGCGCTCCCGCCCGC 213
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 214 ATCCGAGACACCTTCTTCCCGGTTCCACGCGCAGGACTCTCTCCACCGAGGCGCAG 273
QY 81 ProGlyGluProHisProHisValLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 274 CCGGGGAGCGCGCATCCCGACCTCTCTCGACGACGACTCTCTCGCGGGTTTGAGGCGCC 333
```

QY 101 CyValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
Db 334 TGGTTCGAGACATCAAGATCGGGCCATCACTGTCGCCACCGAGTTCCGGGAGCCCTAC 393  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 394 ATCGCCAAAGTACCTCGCCAAAGACCGCGGACCCAGCGAGTTCGTCTCGATTCGCGTC 453  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 454 TTGC--GTCCGAGTCTCGCCCGAGGCGCGGTGTGCGGACGAGCGCCGGAGGTG 511  
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSer-ValAl 180  
Db 512 AAGGCTATGACACCGCTCGGCTCCCGGCTCCCGGCTCTCGGCGCTACGTGTATCCGTTGC 571  
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyLysGlyGlyValLe 200  
Db 572 CGACGAGGGGATGACTGCGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTT 631  
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
Db 632 GTACAGCTGCGGAGCTCAAGGCATGGTTGGAGGAGCACTCTGTTCACCTTCTACTC 691  
QY 220 xAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240  
Db 692 GGGCTCGATTCTCTGGCTATGATGCTGTGACGTCCGACAGCGAGGTGGGGGTGG 751  
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
Db 752 GGTACAGTGAAGTGGTGGACTTTCGCCATGTGGCGGAGGTGTATGGGTGATGACCA 811  
QY 260 sLenPheLeuGlyGlyLeuCysSerLeuLysPheValSerAspIleValProGluTh 280  
Db 812 CAACCTTCCTGGCGAGCTCTGTAGTGTATCAAGTTCGTTTCTGACATTGTTCCAGAGAC 871  
QY 280 rProHisThrGlnProLeuGlyProSer 289  
Db 872 TCCTTAGACGACGCCCTTGGTCTCTCT 899

## RESULT 3

CG453578 818 bp DNA linear GSS 17-SBP-2003  
LOCUS CG9AD40TV ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0849H08,  
DEFINITION genomic survey sequence.

CG453578  
VERSION CG453578.1 GI:34838578

KEYWORDS GSS.

SOURCE Zea mays

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 818)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OG9AD40TH

Contact: Cathy Whitelaw

## TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..818

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

## FEATURES

Source

## RESULT 4

CC344824/c

LOCUS

DEFINITION

CGIAQ20TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0359C15,

genomic survey sequence.

CC344824

VERSION

CC344824.1 GI:30814230

KEYWORDS GSS.

/db\_xref="taxon:4577"  
/clone="ZMMBMA0849H08"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,74e-122 Length: 818  
Score: 1256.00 Matches: 239  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.58% Indels: 0  
DB: 9 Gaps: 0

US-10-042-894A-8 (1-289) x CG453578 (1-818)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 96 ATGTCGACCTCCACCGCGGAGCACCAGTCCGCGCACCGCGCTCCGCCAGAG 155  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 156 CTGGGCGCCCTCATCGACGGCTCCGGCCCTTCTTCAAGCGCTCCAGCGCGGACCG 215  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60  
Db 216 GGGGAGACAGAGTGGCTTCTATAGGGGCTTCTCCGCCACCGCGCGTCCGGCGCG 275  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 276 ATCCGAGACACCTTCTTCCCGGTTCCACGGCAGCATCTCTCCGCCACCGAGGCGCAG 335  
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 336 CCGGGGAGCCGATCTCCTCCTCGCTTCGACGACCTCTCTCGGGGTTTCAGGCGGCC 395  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
Db 396 TCGGTCCGACAGATCAAGATCCGGCCATCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 455  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 456 ATCGCCAAAGTGGCTCGCCAAAGACCGCGGACCCAGCGAGTTCGTCTCGATTCGCGTC 515  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 516 TCGGCGTCCGAGTCTGTCGGCCCGAGGCGCGTGTGGCGGACCGAGCGCCGAGGTG 575  
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSerValAla 180  
Db 576 AAGGCATCGACACCGCGCGGTCTCCCGCGGTCTCCGGCGCTACGTGTCTCCGTTGCC 635  
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyLysGlyGlyValLeu 200  
Db 636 GACGAGGGGATGGACTGTGCGTCTCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTG 695  
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220  
Db 696 TCACAGCTCGCGAGCTCAAGGCGTGTGAGGAGCAGACTCTGTGTCCACTTCTACTCG 755  
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 239  
Db 756 GCCTCGATTCTTCTGGGCTATGATGCTGTGTCAGTTCGACAGCGCGGAGATGGGGG 812

CC344824

CGIAQ20TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0359C15,

genomic survey sequence.

CC344824

VERSION

CC344824.1 GI:30814230

KEYWORDS GSS.

SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
clade; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
1 (bases 1 to 873)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.  
Contact: Cathy Whitelaw  
Unpublished (2002)  
TIGR

TITLE  
JOURNAL 9712 Medical Center Drive, Rockville, MD 20850, USA  
COMMENT Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source Location/Qualifiers  
1..873  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0359C15"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,11e-121 Length: 873  
Score: 1248.00 Matches: 238  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.05% Indels: 0  
DB: 8 Gaps: 0

US-10-042-894A-8 (1-289) x CC344824 (1-873)

QY 52 SerAlaHisAlaValProAlaArgIleArgAspThrPhePheProArgPheHisGly 71  
DB 872 TCGGCCACCGCGCGTCCGGCCCGCATCCGAGACACCTTCTCCCGCGTTCCACGCG 813

QY 72 ThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAsp 91  
DB 812 ACGCGACTCTCCACCGAGCGCAGCGCGGAGCGCATCTCACCCTCGTCTCTCGAC 753

QY 92 AspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIleValIleGlyAlaIleThr 111  
DB 752 GACCTCTCGCGGGTTTCAGCGCCCTCGTGCAGACATCAAGATCGCGCCATCAGC 693

QY 112 TtpProProSerSerProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThr 131  
DB 692 TGGCCACCGAGTTCGCGGAGCCCTACATCGCCAGTGTCTGCCAAGACCGCGGACC 633

QY 132 ThrSerValLeuLeuGlyPheArgValSerGlyValArgValGlyProGluGlyAla 151  
DB 632 ACGAGGTTCTGTCGGATTCGCGTCTCCGGCGTCCGAGTGTGTCGCCCGCGCGCC 573

QY 152 ValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgVal 171  
DB 572 GTGTGGCGACGAGCGCGCGAGGTGAAGGCCATGGACCGCGCGGCTCGCGCGCGTG 513

QY 172 LeuArgArgTyrValSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 191  
DB 512 CTCGGCGGTACGTGTTCATCCGTTGCCGACGAGGGATGTGATGTGCGCTCGCGCGCGC 453

QY 192 ValTyrGlyGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211  
DB 452 GTGTACGGAGAAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGTGGTTCGAG 393

QY 212 GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaA 231  
DB 332 GAGCAGACTCTCTTCCACTTCTACTCGCGTGCATCTTCTGGGCTATGATGCTGCTGCA 333

QY 232 ValAlaAlaGlyGlyVaspglyGlyValThrValLysLeuValAspPheAlaHisVal 251  
DB 332 GTCCGACGAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTG 273

QY 252 AlaGluGlyAspGlyValIleAspHisPheLeuGlyGlyLeuCysSerLeuIleLys 271  
DB 272 GCCGAGGGTGTGGGGTGATTGACCACCACTTCTCTGGCGGGCTCTGCTCGCTGATCAAG 213

QY 272 PheValSerAspIleValProGluThrProHisThrGlnProLeuGlyProSer 289  
DB 212 TTCGTTTTCACATGTTCCGGAGACTCTCATACGACGACCTTTGGGTCCTTCT 159

RESULT 5  
CC724960

LOCUS CC724960 867 bp DNA linear GSS 23-JUN-2003

DEFINITION OGLAB23TV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0307C22,  
genomic survey sequence.

ACCESSION CC724960

VERSION CC724960.1 GI:32143893

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 867)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.  
Contact: Cathy Whitelaw  
Unpublished (2002)  
Other GSSs: OGLAB23TH  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source Location/Qualifiers  
1..867  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0307C22"  
/clone\_lib="ZM 0.7 1.5\_KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.6e-114 Length: 867  
Score: 1183.00 Matches: 236  
Percent Similarity: 90.15% Conservative: 2  
Best Local Similarity: 89.39% Mismatches: 7  
Query Match: 77.78% Indels: 20  
DB: 9 Gaps: 1

US-10-042-894A-8 (1-289) x CC724960 (1-867)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
DB 131 ATGCCCGACCTCCACCCCGGAGACCACCAAGTGGCGGTCCACCGGCTCCCGCAGCAAG 190

QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 191 CTGGGCCCACTCATCGACGACTCTGGCTCTTCTTACAGCCGCTCCAGCGCGGCGACCGT 250



|            |  |  |              |
|------------|--|--|--------------|
| Qy         | 41   | GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg         | 60           |
| Db         | 251  | GGGGAGCAGCAGGTTCGCCTTCTATGAGCGTTCTCCGCCACCGCGCGTCCCGCCCGC            | 310          |
| Qy         | 61   | IleArgAspThrPhePheProAraPheHisGlyThrArgLeuLeuProThrGluAlaGln         | 80           |
| Db         | 311  | ATCCGAGACACTCTTCTCCCCCGGTTCACGGCACGCACTCCTCCCACCGAGCGCAG             | 370          |
| Qy         | 81   | ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro         | 100          |
| Db         | 371  | CCCGGGAGCCGCATCCGCACCTGCTCTCCAGCACTCTCTCGCGGGTTTGAGCGGCC             | 430          |
| Qy         | 101  | CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr            | 120          |
| Db         | 431  | TGCGTCGCAGACATCAAGATCGGTGCCATCAGTG-----ACCACGAGCGTTCTCTCGGATTCCGCGTC | 495          |
| Qy         | 121  | IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal         | 140          |
| Db         | 466  | -----AACACGAGCGTTCTCTCGGATTCCGCGTC                                   | 495          |
| Qy         | 141  | SerGlyValArgValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal            | 160          |
| Db         | 496  | TCCGGCGTCCGAGTCTCGGCCCCGAGGGCGCGTGTGGCGGACGAGCGCCGAGGGTG             | 555          |
| Qy         | 161  | LysAlaMetAspThrAlaGlyValArgArgValLeuArgTyrValSerSer-Valal            | 180          |
| Db         | 556  | AAGCTATGACATTGTGCGCGTCCGCCGCGTCTCCGGCGCTACGTGTCTCATCCGCTTGC          | 615          |
| Qy         | 180  | aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyVal              | 200          |
| Db         | 616  | CGACGAGGGATGGACTCGCGCTCGCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTT           | 675          |
| Qy         | 200  | uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe         | 220          |
| Db         | 676  | GTCAAGCTGCGCGAGCTCAAAGCGGTGTGTGAGGGGCGAGCTCTGTTCACCTTCTTACTC         | 735          |
| Qy         | 220  | rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyGln        | 240          |
| Db         | 736  | GGCGTCGANTCTCTTGGGCTATGATGCTGCTGCTGCACTCGCAGCAGCGGAGGTGGGGTG         | 795          |
| Qy         | 240  | yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi         | 260          |
| Db         | 796  | GGTAACAGTGAAGCTGTGTGGACTTTGCCCATGTGGCCGAGGTGTATGGGGTGAATTGACCA       | 855          |
| Qy         | 260  | sAsnPhelLeu  | 263          |
| Db         | 856  | CAACTTCTCTG  | 865          |
| RESULT     | 6  |  |              |
| LOCUS      | CG284572   |  |              |
| DEFINITION | OGWIF61TH ZM_0.7_1.5_KB Zea mays genomic clone ZMBEma0584K02,<br>genomic survey sequence.  |  |              |
| ACCESSION  | CG284572   |  |              |
| VERSION    | CG284572.1   | GI:34198786  |              |
| KEYWORDS   | GSS.   |  |              |
| SOURCE     | Zea mays   |  |              |
| ORGANISM   | Zea mays   |  |              |
| REFERENCE  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD<br>clade; Panicoideae; Andropogoneae; Zea.<br>1 (bases 1 to 719)<br>Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,<br>Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,<br>Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.<br>Consortium for Maize Genomics<br>Unpublished (2002) |  |              |
| TITLE      | JOURNAL  | Other_GSSs:  | OGWIF61TV    |
| COMMENT    | Contact:   | Cathy Whitelaw   |              |
|            | TIOR   |  |              |
|            |  | 9712 Medical Center Drive, Rockville, MD 20850,                      | USA          |
|            |  | Tel:   | 301-838-5843 |

Fax: 301-838-0208  
Email: whitelawetigr.org  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1. 719  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMa0584K02"  
/clone\_lib="ZM 0.7 1.5\_KB"  
/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

|                        |           |               |     |  |  |
|------------------------|-----------|---------------|-----|--|--|
| Alignment Scores:      |           |               |     |  |  |
| Pred. No.:             | 4,918-111 | Length:       | 719 |  |  |
| Score:                 | 1149.00   | Matches:      | 230 |  |  |
| Percent Similarity:    | 89.9%     | Conservative: | 2   |  |  |
| Best Local Similarity: | 89.1%     | Mismatches:   | 20  |  |  |
| Query Match:           | 75.5%     | Indels:       | 1   |  |  |
| Gaps:                  | 9         |               |     |  |  |
| DB:                    |           |               |     |  |  |

US-10-042-894A-8 (1-289) x CG284572 (1-719)

[illegible]

```

QY 240 YValThrValLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 668 GGTAACAGTGAAGCTGGTGACTTTGCCCATGTGGCCGAGGGTGATGGGGTG 719

RESULT 7
CA134480 801 bp mRNA linear EST 24-SEP-2003
SCJFRT1061H11.g R1 Saccharum officinarum cDNA clone SCJFRT1061H11
5', mRNA sequence.
CA134480
CA134480.1 GI:35021536
EST.
Saccharum officinarum
Saccharum officinarum
ORGANISM

REFERENCE
1 (bases 1 to 801)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 061 row: H column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .801
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFRT1061H11"
/lab_host="DH10B"
/clone_lib="RT1"
/Note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using Superscript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucet.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:
Pred. No.: 3,24e-108 Length: 801
Score: 1123.00 Matches: 220
Percent Similarity: 89.58% Conservative: 12
Best Local Similarity: 84.94% Mismatches: 24
Query Match: 73.83% Indels: 3
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CA134480 (1-801)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 31 ATGTCCGACTTCGGCCGCGGAGACCAAGTCGCGGCCACGGCCCTCCGCCAACAG 90

QY 21 LeuGlyProLeuLAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 91 CTGGGTCCGCTCATCGACGGCTCTGGCTCTTTCTACAGCGCTCCAGTCCGCGGACCGC 150

```

```

QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 151 GGGGAGACAGAGCTCGCTTCTACGAGCGTTCTCCACCACGCGCGCTCCGGCCGC 210

QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 211 ATCCGGGACACCTTCTCCCGCGGTTTACGGCAGCGACTCTCTCCACACGAGGGCGG 270

QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 271 CCAGGGAGCGCATCCGCACCTCGTCTCGACGACCTCTCGCGGGGTGGAGCGGCC 330

QY 101 CysValAlaAspIleValGlyAlaIleThr-TTPProProSerSer-ProGluProT 120
Db 331 TCGTTCGCGCATCAAGATCGCGCCATCAAGTGGCGCGGAGCTCCGCGAGACCT 390

QY 120 YrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgV 140
Db 391 AGTCTCATGTCCTGGTCAAGGACCGCGGACACGAGCATTTCTGTCGATTCGCG 450

QY 140 alSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluV 160
Db 451 TCTCGGGGTCCGGTCCGTCCGAGGGCGCGTGTGGCGACTGAGCGCCGCGAG 510

QY 160 alLysAlaMetAspThrAlaGlyValArgValLeuArgValSerValSerVal 180
Db 511 TGAAGCCCTTGACACCCGCGCGTCCGCGGCTCCGCGCTACGTTTCATCGGTG 570

QY 180 laAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyVal 200
Db 571 CGACGAGGGATGACCTGCGACCTCGCGCGCGGTGTACGCGGGAAGGGAGTTT 630

QY 200 euSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrS 220
Db 631 TGTCAAAATTGGCGAGCTTAAGCGGTGTTTGAAGAAACAGACTCTGGTCCACCTTACT 690

QY 220 erAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaGlyGly-AspGlyGly 239
Db 691 CGGGGGCGATTTCTCTGGGCTATGAAGCTAGTGGCTGCAACACCGGAAGTAAAGGT 750

QY 240 GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256
Db 751 GGGGTGAAGTAAACTGGGGGCCCTTTGCCATGTGGCCCAAGGGAAGG 801

RESULT 8
CA245555 656 bp mRNA linear EST 25-SEP-2003
SCBGF5081A03.g Saccharum officinarum FL5 Saccharum officinarum
cDNA clone SCBGF5081A03 5', mRNA sequence.
CA245555
CA245555.1 GI:35324301
EST.
Saccharum officinarum
Saccharum officinarum
ORGANISM

REFERENCE
1 (bases 1 to 656)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br

```

Plate: 081 row: A column: 03  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers

## FEATURES

source

1..656  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCBGL5081A03"  
/lab\_host="DH10B"  
/clone\_lib="Saccharum officinarum FL5"  
/notes="Organ: Developed inflorescence (20cm-long) without rachis; Vector: pSport1; Site 1: Sali; Site 2: NotI; An unidirectional cDNA library generated from [Developed] inflorescence (20cm-long) without rachis). cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of source of RNA and library construction can be obtained at <http://sucet.lad.ic.unicamp.br/public>"

## ORIGIN

Alignment Scores:  
Pred. No.: 6,53e-99 Length: 656  
Score: 1034.00 Matches: 196  
Percent Similarity: 97.60% Conservative: 7  
Best Local Similarity: 94.23% Mismatches: 4  
Query Match: 67.98% Indels: 1  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CA245555 (1-656)

Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db |||||  
Qy 33 ATGTCCGACCTCCGCCCGCGGACCAAGTCCCGGCCACCGCGCTCCGCCAACAA 92  
Db |||||  
Qy 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db |||||  
Qy 93 CTGGGTCCGCTCATCGACGGCTTGGGCTCTTTCTACAAGCGCTCTCCAGGTCCGGGACCGC 152  
Db |||||  
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
Db |||||  
Qy 153 GGGGAGCAGCGTCCGCTTCTACAGCGCTTCTCCACCGCGCGCGTCCCGGCCGC 212  
Db |||||  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db |||||  
Qy 213 ATCCGGACACCTTCTTCCCGCGGTTTTCACGGCACCGGACTCTCTCCACCGAGCGCGG 272  
Db |||||  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db |||||  
Qy 273 CCAGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCTCGCGGGCTGGAGCGGCC 332  
Db |||||  
Qy 101 CysValAlaAspIleLysIleGlyValAlaIleThrTrpProProSerSerProGluProTyr 120  
Db |||||  
Qy 333 TGGTCCCGGACATCAGATCGGCCATCAGTGGCGCGGAGCTCGCGGAGCCCTAC 392  
Db |||||  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db |||||  
Qy 393 GTCCCAAGTGCTTCGCCAAGGACCGGGGACCAACGAGCATTTCTGCTCGGATTCGGGTC 452  
Db |||||  
Qy 141 SerGlyValArgValValGlyProGluGlyValAlaValTrpArgThrGluArgProGluVal 160  
Db |||||  
Qy 453 TCCGGCTCCGGGTCGTGGTCCGAGGGCGCGCTGTGGCGGACTGAGCGGCCGGAGGTG 512  
Db |||||  
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180  
Db |||||  
Qy 513 AAGCCCTGGACACCGCGCGGCTCCGCGGTGTCTCCGGCGCTACGTTTCATCGTTGCC 572  
Db |||||  
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGly-LysGlyGlyValLe 200  
Db |||||  
Qy 573 GACGAGGGGATGGACTCGCACTCGCGCGCGGTGTACGCGCGCAAAAGGGGGAGTCTT 632  
Db |||||

Qy 200 userGlnLeuArgGluLeuLys 207  
Db |||||  
633 GTCACAGCTGGCGAGCTTAAG 654

## RESULT 9

CA130685

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

CA130685

VERSION

CA130685.1

KEYWORDS

SACCHARUM OFFICINARUM

ORGANISM

SACCHARUM OFFICINARUM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

REFERENCE

1 (bases 1 to 722)

AUTHORS

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

TITLE

The libraries that made SUCST

JOURNAL

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bccc.bccp.br>

Plate: 004 row: G column: 05

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1..722

/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCCCT1004G05"

/lab\_host="DH10B"

/clone\_lib="RT1"

/notes="Organ: Root tips (0.3cm-long) from adult plants;

Vector: pSport1; Site 1: Sali; Site 2: NotI; An

unidirectional cDNA library generated from [Root tips

(0.3cm-long) from adult plants]. cDNA was prepared from

polyA+ mRNA using SuperScript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

<http://sucet.lad.ic.unicamp.br/public>"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,45e-94 Length: 722  
Score: 993.50 Matches: 202  
Percent Similarity: 92.07% Conservative: 7  
Best Local Similarity: 88.99% Mismatches: 17  
Query Match: 65.32% Indels: 4  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x CA130685 (1-722)

Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20

Db |||||

31 ATGTCCGACCTCGC-CGCGCGGAGCACCAGTCCCGGCCACCGCGCTCCGCCAACAA 89

Qy 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40

Db |||||

90 CTGGGTCCGCTCATCGACGGCTCTGGCTCTTCTACAAGCGCTCTCCAGGTCCGGGACCGC 149

Db |||||

```
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 150 GGGAGCACGAGCTCGCTTCTACGAGGCGTTCTCCACCCACGCCCGCTCCGGCCGC 209
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 210 ATCCGGACACCTTCTTCCCGCGGTTTACGAGCAGCGACTCTCCACCGAGGGCGG 269
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 270 CCAGGGAGCGCATCCGACCTCGTCTCGAGACCTCTCCGGGGCTGGAGGGCGCC 329
Qy 101 CysValAlaAspIleYsIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 330 TCGGTGCGGACATCAAGATCGGCGCATCAGCTGGCGCGAGCTCGCGGAGCCCTAC 389
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 390 GTGCCCAAGTGCCTCGCCAGGACCGCGGACCAAGGATTCGTCGATTCGCGTTC 449
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 450 TCCGGGTCCGGTCTCGGTCGCCAGGGCGCGTGTGGCGGACTGAGCGCCCGGAGGTG 509
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla 180
Db 510 AAGGCCCTGGACACCGCGCGCTCCGCGCTCTCCGGCGCTTACGTTTCATCCGTTGCC 569
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaVal 200
Db 570 GACGAGGGGATGACTCGGCACCTCGCGCGCGGGGAAACCGCGCGGAGGAGTCTT 629
Qy 200 uSerGlnLeuArgGluLeuYsAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220
Db 630 GTCAAGCTCCCGAGCTTAAGCGTGGTTCGAGGA-CAAACTCTGGATCACTT-TACTC 687
Qy 220 rAlaSerIleLeuLeuGly 226
Db 688 GGGGCGGATCTTTCGGGC 706
```

```
RESULT 10
CG776236 591 bp DNA linear GSS 29-OCT-2003
LOCUS 1123005E08.x1 1123 - RescueMu Grid L Zea mays genomic, genomic
DEFINITION survey sequence.
```

```
ACCESSION CG776236
VERSION 1.1
KEYWORDS GSS.
SOURCE Zea mays
```

## ORGANISM

Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 591)

Walbot, V.

Maize genomic sequences found using engineered RescueMu transposon  
Unpublished (2001)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1123005 column: 3

Class: transposon-tagged.

Location/Qualifiers

1..591

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A188/B73/K55"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

## FEATURES

source

```
/dev_stage="adult"  
/lab_host="DH10B"  
/clone_lib="1123 - RescueMu Grid L"  
/note="Organ: leaf; Vector: RescueMu (engineered from  
pBluescript backbone); Site 1: BamHI; Site 2: BglII;  
RescueMu is a 4.9 kb, modified maize Mu transposon  
designed to allow plasmid rescue from total genomic DNA.  
Mu elements insert preferentially into transcription  
units. For more information on RescueMu, go to the web  
site 'www.zmdd.iastate.edu' and follow the links for  
'RescueMu.' Grid L was grown in Molokai in 2001. DNA was  
extracted from leaf strips, double digested using BamHI  
and BglII, and ligated to form circular plasmids. DH10B  
cells were transformed and then screened on LB plates with  
ampicillin."
```

## ORIGIN

```
Alignment Scores:  
Pred. No.: 1.58e-94 Length: 591  
Score: 992.00 Matches: 187  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.47% Mismatches: 0  
Query Match: 65.22% Indels: 0  
DB: 9 Gaps: 0
```

US-10-042-894A-8 (1-289) x CG776236 (1-591)

```
Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerHisLys 20  
Db 28 ATGTCCGACCTCCACCGCGGAGCACCAGTCCGCGCCACCGCGCTCCGCCAGCAAG 87  
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 88 CTGGGCGCGCTCATCGACGCTCCGGCTCTTCTTAAAGCCGCTCTCAGCGCGCGCGCT 147  
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60  
Db 148 GGGAGACGAGGTGCTCTTCTATGAGGCTTCTCCGCCACCGCGCTCCGCCCGCC 207  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 208 ATCCGAGACACCTTCTCCCGGTTCCACGCGACGACTCTCCGCCACCGAGGCGCAG 267  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
Db 268 CCGGGAGCGCGATCTCCTCTCTCGACGACCTCTCTCGGGGTTTCAGGCGCCC 327  
Qy 101 CysValAlaAspIleLeuIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 328 TCGGTGCGGACATCAAGATCGGCGCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 387  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 388 ATCGCCAAAGTGTCTCGCCAAAGACCGCGGACCAAGGAGGTTCTCTCGGATTCGCGCTC 447  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 448 TCGGCGTCCGAGTCTGTGCGCGCGCGCGCTGTGCGGACCGAGGCGCGCGGAGGTG 507  
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla 180  
Db 508 AAGGCCATGGACACCGCGCGCTCCGCCGCTGTCTCCGCGCTCTCCGCGCTCCGCTGCC 567  
Qy 181 AspGluGlyMetAspCysAlaLeu 188  
Db 568 GACCAGGGGATGAGCTGTGCGCTC 591
```

## RESULT 11

CA202412

LOCUS

DEFINITION

ACCESSION

CA202412 671 bp mRNA linear EST 25-SEP-2003  
SCRLFL1009H02.g FL1 Saccharum officinarum cDNA clone SCRLFL1009H02  
5', mRNA sequence.  
CA202412

```

VERSION          CA202412.1  GI:35237651
KEYWORDS          EST.
SOURCE            Saccharum officinarum
ORGANISM          Saccharum officinarum

REFERENCE
AUTHORS           Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE             The libraries that made SUCBST
JOURNAL           Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT           Contact: Arruda P
                  Centro de Biologia Molecular e Engenhariaia Genetica
                  Universidade Estadual de Campinas
                  Caixa Postal 6010, 13083-970, Campinas SP, Brazil
                  Tel: 55 19 3788 1137
                  Fax: 55 19 3788 1089
                  Email: parruda@unicamp.br
                  Clone distribution: clone distribution information can be found
                  through the Brazilian Clone Collection Center (BCCC) at
                  http://www.bcccenter.fcav.unesp.br
                  Plate: 009 row: H column: 02
                  Seq primer: T7 Promoter Primer.
FEATURES          Location/Qualifiers
                  1..671
                    /organism="Saccharum officinarum"
                    /mol_type="mRNA"
                    /db_xref="taxon:4547"
                    /clone="SCRLFL1009H02"
                    /lab_host="DH10B"
                    /clone_lib="FL1"
                    /note="Organ: Inflorescence at beginning of development
                    (1cm-long); Vector: pSport1; Site_1: SalI; Site_2: NotI;
                    An unidirectional cDNA library generated from
                    [Inflorescence at beginning of development (1cm-long)].
                    cDNA was prepared from polyA+ mRNA using Superscript
                    Plasmid System kit (Invitrogen). The double-strand cDNAs
                    were fractionated in a sepharose CL-2B 40cm-columns and
                    fragments sizing between 0.8 and 1.5 Kb were
                    directionally cloned into the vector. Details
                    of source of RNA and library construction can be obtained at
                    http://sucst.lad.ic.unicamp.br/public"
ORIGIN
Alignment Scores:
Pred. No.:        6,4e-94      Length:      671
Score:            987,00      Matches:    192
Percent Similarity: 91,93%    Conservative: 13
Best Local Similarity: 86,10%  Mismatches: 17
Query Match:      64,89%      Indels:     2
DB:               6           Gaps:       0

US-10-042-894A-8 (1-289) x CA202412 (1-671)
Qy  41  GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db  2   GGGAGACAGAGCTCGCTCTTACGAGCGGTCTCCAACACACGCGCGCTCCCGCCGC 61
Qy  61  IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db  62  ATCGGGACACCTTCTCCCGCGGTTCACGGACGCGATCTCTCCACCGAGCGCG 121
Qy  81  ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db  122  CCAGGGAGCGCATCGACCTCGTCTCGACGACCTCTCGCGCGGCTCGAGCGCGCC 181
Qy  101  CysValAlaAspIleValIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db  182  TGCCTCGCCGACATCAAGATCGGCGCCATCAGTGGCGCGGAGCTCGCGGAGCCCTAC 241
Qy  121  IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140

```

---

```

Db  242  GTCCCAAGTCTCGCCAAAGGACCGCGGACCCACAGAGATTCTGTCTGGATTCCGCGTC 301
Qy  141  SerGlyValArgValValGlyProGluGlyAlaValTirArgThrGluArgProGluVal 160
Db  302  TCGGGCTCGGGTCTCGTCCGAGGGCGCGGTGTGGCGACTGAGCGCCCGGAGGTG 361
Qy  161  LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
Db  362  AAGGCCCTGGACACCGCGCGTCCGCGCGTCTCGCGCGTACGTTTCATCCGTTGCC 421
Qy  181  AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200
Db  422  GACGAAGGATGAGTACGCGACTCGCGCGCGGTGTACGCGCGGCGGAGGAGGAGTTC 481
Qy  201  SerGlnLeuArgGluLeuLysAlaTirPheGluGluGlnThrLeuPheHisPheTyrSer 220
Db  482  TCACAGCTCGCGAGCTTAAGCGGTGTTCGAGGAGCAGACTCTGGTCCACTTCTACTCG 541
Qy  221  AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db  542  GCGGTCAATCTTTTGGGCTATGATGCTAGTGCAAGTGCAGCAAGCGAGGTGAAGGTGG 601
Qy  241  ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp-GlyValIleAspHi 260
Db  602  TTGAAGGTGAA-CTTGTGACTTTTCCCATTTTGTCCAAAGTTGTAGTAGGGGAATGAACCA 660
Qy  260  'eaePhe 262
Db  661  AATTTTC 667

RESULT 12
AY109355      3374 bp      mRNA      linear      HTC 17-OCT-2002
LOCUS        Zea mays Cl49_2 mRNA sequence.
DEFINITION   AY109355
ACCESSION    AY109355.1  GI:21213015
VERSION      HTC.
KEYWORDS      Zea mays
SOURCE       Zea mays
ORGANISM     Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 3374)
AUTHORS      Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
              Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
              Maize Mapping Project/DuPont Consensus Sequences for Design of
              Overgo Probes
              Unpublished (2002)
              2 (bases 1 to 3374)
              Coe,E.H.
              Direct Submission
              Submitted (25-APR-2002) Maize Mapping Project, University of
              Missouri, Columbia, MO 65211, USA
              If you are interested in getting corresponding physical clones,
              these are publicly available from ZmDB and may be found by BLAST
              searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
              www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
              maize cDNA sequences is either Virginia Walbot, Stanford or Pat
              Schnable, Iowa State, then clones may be requested from ZmDB:
              www.zmdb.iastate.edu.
FEATURES      Location/Qualifiers
              1..3374
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="MaizeDB:632179"
                /db_xref="taxon:4577"
                /clone_lib="Maize Mapping Project/DuPont Consensus
                Library"
                /note="this sequence is part of a project of EST
                assemblies resulting from the application of public
                contigs to seed Dupont contigs; this resource was
                assembled by Dupont as part of a collaboration for the
                overgo addressing of BACs in conjunction with the Maize

```

## Mapping Project"

## ORIGIN

Alignment Scores: 2.63e-92 Length: 3374  
Pred. No.: 981.00 Matches: 205  
Score: 981.00  
Percent Similarity: 71.88% Conservative: 2  
Best Local Similarity: 71.18% Mismatches: 62  
Query Match: 64.50% Indels: 20  
DB: 3 Gaps: 1

US-10-042-894A-8 (1-289) x AY109355 (1-3374)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisLeuLys 20  
DB 72 ATGCCCGACTCCACCGCGGAGACCAAGTCGCGGTCACCGCCCTCCGCGACCAAG 131  
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 132 CTGGGCCCACTCATCGACGACTCTGGCCTCTTCTACAAGCGCTCCAGGCGCGGACCGT 191  
QY 41 GlyLHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 192 GGGAGCAGCGAGTCCGCTTCTATGAGCGGNNNNNNNNNNNNNNNNNNNNNNNNNN 251  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 252 NTCGAGACACCTTCTNNNNNGTTCCAGGACGCGACTCTCCCAACGAGGCGAG 311  
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100  
DB 312 CCGGGGAGCGCATCCGACCTCGCTCTCGACGACTCTCTCGCGGGTTTGAGGCGCC 371  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
DB 372 TGGTGCAGACATCAAGATCGGTGCCATCACGTG----- 406  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 407 -----ACCACGAGCGTCTGCTCGGATCCGCGTC 436  
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
DB 437 TCGCGGTCGAGTCTCGCGCCGAGGCGCGTGNNNNNNNNNNNNNNNNNNNNNNNNG 496  
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
DB 497 AAGGCTATGGACATTCTCGCGCGTCCGCGCTGCTCCGCGCTACGTGTCTCCGCTGC 556  
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeu 200  
DB 557 CGACGAGGGATGGATCGCGCTCCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTT 616  
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
DB 617 GTCACAGCTGCGGAGCTCAAGCGGTGTTCGAGGGGCGAGACTNNNNNNNNNNNNNN 676  
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGly 240  
DB 677 NNN 736  
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAsphi 260  
DB 737 NNN 796  
QY 260 sAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
DB 797 CAACTTCCTGGGCGGCTCTGCTAGCTGATCAAGTTTGTCTGACATTGTTCCAGAGAC 856  
QY 280 rProHisThrGlnProLeuGly 287  
DB 857 TCTTCAGACGCGCCCTTGGG 878

## RESULT 13

## CC724950/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## Location/Qualifiers

## 1..722

## /organism="Zea mays"

## /mol\_type="genomic DNA"

## /strain="B73"

## /db\_xref="caxon:4577"

## /clone="ZMMBMA0307C22"

## /clone\_lib="ZM 0.7 1.5\_KB"

## /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb"

## methylation filtered genomic DNA library"

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Score:

## Percent Similarity:

## Best Local Similarity:

## Query Match:

## DB:

## US-10-042-894A-8 (1-289) x CC724950 (1-722)

## QY 66 PheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluProHis 85

## DB 721 TTCCTCCCGGTTCCACGGCAGCGACTCTCCCAACGAGCGCGAGCGGAGCGCAT 662

## QY 86 ProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIle 105

## DB 661 CGCACCTCTGCTCTCGACGACTCTCTCGCGGGTTTGAGGCGCCCTGCTCGACACATC 602

## QY 106 LysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLysCysLeu 125

## DB 601 AAGATCGGTGCCATCACGTG----- 582

## QY 126 AlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145

## DB 581 -----ACCACGAGCGTCTGCTCGGATTCGCGCTCCGCGGCTCCGAGTC 537

## QY 146 ValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAlaMetAspThr 165

## DB 536 GTCGCGCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGAGGTGTAAGCTATGACATT 477

## QY 166 AlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAlaAspGluGlyMetAs 185

## DB 476 GTCGCGCTCCGCGGCTGCTCCGCGGCTACGTGTCTATCCGCTTCCGAGGGGATGA 417



http://sucest.lad.ic.unicamp.br/public"

|            |     |   |     |
|------------|-----|---|-----|
| QY         | 185 | pCyEAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgG1   | 205 |
| Db         | 416 | CTGGCGCTCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTTGTACAGCTGCGCGA     | 357 |
| QY         | 205 | uLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLe   | 225 |
| Db         | 356 | GCTCAGGCGTGTTCGAGGGGCGAGCTCTGTTCACCTTCTACTCGGCGTGGATTCTCT   | 297 |
| QY         | 225 | uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGlyValThrValLysLe   | 245 |
| Db         | 296 | GGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT     | 237 |
| QY         | 245 | uValAspPheAlaHisValAlaGluGlyAspGlyValLleAspHisPheLeuGlyG1   | 265 |
| Db         | 236 | GGTGGAGCTTTGGCCATGTGCGGAGGTGTGGGGTGTGGGGTGTGGGGTGTGGGG      | 177 |
| QY         | 265 | YLeuCySerLeuIleLysPheValSerAspIleValProGluThrProHisThrGlnPr | 285 |
| Db         | 176 | GCTCTGCTAGCTGATCAAGTTTGTCTGACATTGTTCCAGAGACTCTCTCAGACGCC    | 117 |
| QY         | 285 | oLeuGlyProSer 289   |     |
| Db         | 116 | TTTGGGTCTCTCT 104   |     |
| RESULT 14  |     |   |     |
| CA130686   |     |   |     |
| LOCUS      |     |   |     |
| DEFINITION |     |   |     |
| ACCESSION  |     |   |     |
| VERSION    |     |   |     |
| KEYWORDS   |     |   |     |
| SOURCE     |     |   |     |
| ORGANISM   |     |   |     |
| REFERENCE  |     |   |     |
| AUTHORS    |     |   |     |
| TITLE      |     |   |     |
| JOURNAL    |     |   |     |
| COMMENT    |     |   |     |
| FEATURES   |     |   |     |
| source     |     |   |     |





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 05:35:00 ; Search time 2130 Seconds  
(without alignments)  
842.244 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10042894/runat\_15062005\_111952\_8386/app\_query.fasta\_1.455  
-DB=Published Applications NA -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODES=LOCAL -OUTFM=ptp -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10042894 @CGN 1 1 480 @runat\_15062005\_111952\_8386  
-NCPU=6 -ICPU=3 -NO MAP -LARG QUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID                   | Description       |
|------------|--------|-------------|--------|----|----------------------|-------------------|
| 1          | 1521   | 100.0       | 1344   | 14 | US-10-042-894A-7     | Sequence 7, Appl  |
| 2          | 1485   | 97.6        | 923    | 14 | US-10-042-894A-5     | Sequence 5, Appl  |
| 3          | 1484   | 97.6        | 1426   | 20 | US-10-425-115-74438  | Sequence 74438, A |
| 4          | 1435   | 94.3        | 923    | 14 | US-10-042-894A-3     | Sequence 3, Appl  |
| 5          | 1406.5 | 92.5        | 1169   | 14 | US-10-042-894A-1     | Sequence 1, Appl  |
| 6          | 1307   | 85.9        | 3416   | 14 | US-10-042-894A-20    | Sequence 20, Appl |
| 7          | 1233   | 81.1        | 899    | 14 | US-10-042-894A-15    | Sequence 15, Appl |
| 8          | 1148.5 | 75.5        | 1108   | 19 | US-10-437-963-4650   | Sequence 4650, Ap |
| 9          | 847.5  | 55.7        | 696    | 19 | US-10-437-963-4649   | Sequence 4649, Ap |
| 10         | 808.5  | 53.2        | 643    | 14 | US-10-042-894A-17    | Sequence 17, Appl |
| 11         | 742.5  | 48.8        | 1020   | 14 | US-10-042-894A-13    | Sequence 13, Appl |
| 12         | 722.5  | 47.5        | 1195   | 14 | US-10-042-894A-11    | Sequence 11, Appl |
| 13         | 644.5  | 42.4        | 1105   | 14 | US-10-042-894A-9     | Sequence 9, Appl  |
| 14         | 644.5  | 42.4        | 1345   | 18 | US-10-424-599-106467 | Sequence 106467,  |
| 15         | 609    | 40.0        | 776    | 20 | US-10-425-115-74436  | Sequence 74436, A |
| 16         | 464    | 30.5        | 519    | 14 | US-10-042-894A-18    | Sequence 18, Appl |
| 17         | 329    | 21.6        | 353    | 14 | US-10-042-894A-19    | Sequence 19, Appl |
| 18         | 306.5  | 20.2        | 484    | 10 | US-09-770-961-829    | Sequence 829, App |
| 19         | 273.5  | 18.0        | 464    | 9  | US-09-770-444-340    | Sequence 340, App |
| 20         | 253    | 16.6        | 295    | 20 | US-10-425-115-10743  | Sequence 10743, A |
| 21         | 201.5  | 13.2        | 2212   | 20 | US-10-478-146-20     | Sequence 20, Appl |
| 22         | 193    | 12.7        | 1341   | 18 | US-10-641-643-809    | Sequence 809, App |
| 23         | 192    | 12.6        | 2890   | 18 | US-10-415-011-37     | Sequence 37, Appl |
| 24         | 191    | 12.6        | 2608   | 17 | US-10-094-749-1021   | Sequence 1021, Ap |
| 25         | 191    | 12.6        | 2608   | 21 | US-10-887-553A-1159  | Sequence 1159, Ap |
| 26         | 189    | 12.4        | 1737   | 9  | US-09-731-872-69     | Sequence 69, Appl |
| 27         | 189    | 12.4        | 1737   | 10 | US-09-876-997-69     | Sequence 69, Appl |
| 28         | 189    | 12.4        | 1737   | 21 | US-10-643-836-69     | Sequence 69, Appl |
| 29         | 189    | 12.4        | 1737   | 21 | US-10-956-157-2095   | Sequence 2095, Ap |
| 30         | 189    | 12.4        | 1748   | 15 | US-10-037-270-712    | Sequence 712, App |
| 31         | 189    | 12.4        | 1748   | 17 | US-10-117-722-712    | Sequence 712, App |
| 32         | 189    | 12.4        | 1757   | 9  | US-09-731-872-76     | Sequence 76, Appl |
| 33         | 189    | 12.4        | 1757   | 10 | US-09-876-997-76     | Sequence 76, Appl |
| 34         | 189    | 12.4        | 1757   | 21 | US-10-643-836-76     | Sequence 76, Appl |
| 35         | 189    | 12.4        | 1791   | 14 | US-10-198-846-10945  | Sequence 10945, A |
| 36         | 189    | 12.4        | 1990   | 9  | US-09-731-872-174    | Sequence 174, App |
| 37         | 189    | 12.4        | 1990   | 20 | US-09-876-997-174    | Sequence 174, App |
| 38         | 189    | 12.4        | 1990   | 21 | US-10-643-836-174    | Sequence 174, App |
| 39         | 182    | 12.0        | 1863   | 17 | US-10-120-988-103    | Sequence 103, App |
| 40         | 178.5  | 11.7        | 1893   | 17 | US-10-120-988-102    | Sequence 102, App |
| 41         | 168    | 11.0        | 4461   | 19 | US-10-618-941-57     | Sequence 57, Appl |
| 42         | 152.5  | 10.0        | 546    | 19 | US-10-437-963-5687   | Sequence 5687, Ap |
| 43         | 145.5  | 9.6         | 1782   | 9  | US-09-954-456-87     | Sequence 87, Appl |
| 44         | 145.5  | 9.6         | 1782   | 9  | US-09-954-456-722    | Sequence 722, App |
| 45         | 145.5  | 9.6         | 1782   | 14 | US-10-175-523-196    | Sequence 196, App |

#### ALIGNMENTS

##### RESULT 1

US-10-042-894A-7  
; Sequence 7, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shit, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafaleki, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase  
; FILE OF INVENTION: 1286  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465

```
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(921)
US-10-042-894A-7

Alignment Scores:
Pred. No.: 4,71e-172 Length: 1344
Score: 1521.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-7 (1-1344)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
DB 52 ATGTCCGACCTCCACCCGCGGAGCACCAGTCCGCCGCCACCGCGCTCCGCCAGCAAG 111
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 112 CTGGGCGCGCTCATCGACGCGCTCCGCGCTCTTCTACAAGCGCGCTCCAGCGCGCGACCGT 171
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 172 GGGGACACAGAGTCCGCTCTATGAGGCGTCTCCGCCACCGCGCGCTCCGCCCGCC 231
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 232 ATCCGAGACACCTTCTCCCGCGTTCACGGCACCGGACTCTCCACCGCGCGCGAG 291
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
DB 292 CCGCGGGAGCGCATCTCACCTCGCTCTCGACGACTCTCTCGCGGGTTTCAGCGCGCC 351
QY 101 CysValAlaAspIleValIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 352 TGGTGCAGACATCAAGATCGCGCGCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 411
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 412 ATGCCCAAGTGCCTGCCAAGGACCGCGGACCACGAGCGTCTCTCGGATTCGCGCTC 471
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
DB 472 TCGGCGCTCGAGTCTCGCGCCCGAGGGCGCGTGTGGCGGACGAGCGCGCGAGGTG 531
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
DB 532 AAGGCGATGACACCGCGCGCTCCCGCGCTCTCCGCGCTACGTGTCATCCGTGCG 591
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeu 200
DB 592 GACGAGGGGATGAGTGTGCGCTCGCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTG 651
QY 201 SerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSer 220
DB 652 TCACAGCTGCGGAGGTCAGGCGTGGTTCGAGGAGCAGACTCTGTTCACACTTCTACTCG 711
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGly 240
DB 712 GCCTCATCTCTCGGCTATGATGCTGCTGCAAGTGCACAGCAGCGGAGATGGGGTGGG 771
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 772 GTGACGGTGAAGCTGTGGACTTTTGCCCATGTGGCCGAGGGTGATGGGGTGATTGACCAC 831

; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(921)
US-10-042-894A-5

Alignment Scores:
Pred. No.: 6e-168 Length: 923
Score: 1485.00 Matches: 283
Percent Similarity: 98.27% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 5
Query Match: 97.63% Indels: 0
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-5 (1-923)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
DB 53 ATGCCCGACCTCCACCCGCGGAGCACCAGTCCGCGGTCCACCGCGCTCCGCCAGCAAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CCGGCGCGCGCTCATCGACGCGTCCGCGCTCTTCTACAAGCGCGCTCCAGCGCGCGACCGT 172
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 173 GGGGAGCAGAGTTCGCTTCTATGAGGCGTTCCTCCGCCACCGCGCGCTCCGCCCGCC 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTCCCGCGTTCACGCGCAGCGACTCTCCCCACCGAGGGCGAG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGCGGGAGCGCGCATCCGACCTCGCTCTCGACGACCTCTCTCCGCGGATTTGAGCGCGCC 352
QY 101 CysValAlaAspIleValIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 353 TCGGTCGAGACATCAAGATCGCGCGCATCAGTGGCCACCGAGTTCGCGGAGCGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
```

```
Db 413 ATGCCCAAGTGCCTGCCATGACCGCGGACACAGAGCGTTCTGCTCGGATTCGCGTC 472
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrrArgThrGluArgProGluVal 160
Db 473 TCCGGGTCCGAGTCGTCGCCCGGAGGCGCGTGTGGCGGACGGAGCGCCCGGAGGTG 532
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
Db 533 AAGGCCATGGACACCGCGCGGTCCGCGGTGCTCGGGGCTACGTGTCTATCGTTGCC 592
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyGlyGlyValLeu 200
Db 593 GACGAGGGATGACACTGTGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTTG 652
Qy 201 SerGlnLeuArgGluLeuYsAlaTrrPheGluGlnThrLeuPheHisPheTyrSer 220
Db 653 TCACAGCTGCGGAGCTCAAGGCGGTGTTCGAGAGCAGACTCTGTTCACATTCCTACTCG 712
Qy 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db 713 GCCTCGATTTCTTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
Qy 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
Db 773 GTGACGGTGAAGCTGGTGGACTTTGCCATGTGGCCGAGGGTGATGGGGTATTGACCCAC 832
Qy 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
Db 833 AACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTTCTGTCATTTGTTCCAGAGACT 892
Qy 281 ProHisThrGlnProLeuGlyProSer 289
Db 893 CCTCAGACGCGCTTTGGGTCTCTTCT 919
```

## RESULT 3

```
US-10-425-115-74438
; Sequence 74438, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 74438
; LENGTH: 1426
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167886C.1
US-10-425-115-74438
```

## Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 1,39e-167 | Length:       | 1426 |
| Score:                 | 1484.00   | Matches:      | 286  |
| Percent Similarity:    | 98.96%    | Conservative: | 0    |
| Best Local Similarity: | 98.96%    | Mismatches:   | 3    |
| Query Match:           | 97.57%    | Indels:       | 1    |
| DB:                    | 20        | Gaps:         | 0    |

US-10-042-894A-8 (1-289) x US-10-425-115-74438 (1-1426)

```
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 120 ATGTCGACCTCCACCGCGCGGAGCACCAAGTCGCGCGCCACCGCGCTCCGCGAGG 179
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
```

```
Db 180 CTGGCGCCGCTCATCGACGCGCTCCGCGCTTCTTACAAAGCGCTCCAGCGCGCGACCGT 239
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 240 GGGGAGCAGAGTCGCTTCCTATGAGGCGTTCTCCGCCACCGCCGCTCCCGGCGCCG 299
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 300 ATCCGAGACACCTTCTTCCCGGTTCCACGCGCAGCGACTCTCTCCCCACCGAGGCGCAG 359
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db 360 CCGGGGAGCGCATCTCCTACCTCGTCTCGACAGCTCTCTCCGCGGGTTTCAGGCGCCC 419
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrrProSerSerProGluProTyr 120
Db 420 TCGCTCGCAGACATCAAGATCGCGCCATCACTGTGCCACCGAGTTTCGCGGAGCCCTAC 479
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 480 ATCCCAAGTCCTCGCCAAAGACCGCGGACCCAGCGGTTCTGCTCGGATTCGCGTC 539
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrrArgThrGluArgProGluVal 160
Db 540 TCCGCGCTCCGAGTCGTCGCGCCCGAGGCGCGCTGTGGCGGACGGAGCGCCGAGGTG 599
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
Db 600 AAGCCATCGACACCGCGCGGTCTCCGCGCTACGTGTCTCGGAGTTCATTCGTTGCC 659
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyLysGlyGlyValLeu 200
Db 660 GACGAGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTTG 719
Qy 201 SerGlnLeuArgGluLeuLysAlaTrrPheGluGlnThrLeuPheHisPheTyrSer 220
Db 720 TCACAGCTCGCGAGCTCAAGGCGTGGTTTCGAGGAGCAGACTCTGTTCACATTCCTACTCG 779
Qy 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db 780 GCGTCGATTTCTTGGGCTATGATGCTGTTGAGTCCA-CGACCGGAGATGGGGTGGG 838
Qy 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
Db 839 GTGACCGTGAAGCTGTGGACTTTGCCCATGTGCGCGAGGGTGTATGGGGTATTGACCCAC 898
Qy 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
Db 899 AACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTTCTGTCATTTGTTCCGAGACT 958
Qy 281 ProHisThrGlnProLeuGlyProSer 289
Db 959 CCTCATACGACGCTTTGGGTCTCTTCT 985
```

## RESULT 4

```
US-10-042-894A-3
; Sequence 3, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1e1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(736)
US-10-042-894A-3

Alignment Scores:
Pred. No.: 5,97e-162 Length: 923
Score: 1435.00 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 94.35% Indels: 1
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-3 (1-923)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 53 ATCCCGAGCTTCACCGCGGAGCACCAGTCGCGGTCACCGCGCTCCCGCCAGCAAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 113 CCGGGCCCGCTCATCGAGCGCTCCGGCTCTTCTACAAGCGCTCCAGCGCGGACCGT 172
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 173 GGGGAGCAGAGGTGCGCTTTCTATAGGGGTTCTCCGCCACGCGCGCTCCCGGCCGC 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 233 ATCCGAGACACCTTCTCCCGCGTTCCAGCGCAGCGACTCTCTCCACCGAGGCGCAG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 293 CCGGGGAGCGCATCGCACCTCGCTCTCGAGACCTCTCGCGGGATTGAGGGCGCC 352
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 353 TGGCTCGCAGACATCAAGATCGCGGCATCATCGTGCCACCGAGTTCGCGGAGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgVal 140
Db 413 ATGCCAAGTGCCTCGCCATGACCGCGGAGCACCAGCGTTCGTCTCGATTCGCGGTC 472
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 473 TCCGGGCTCCGAGTCTGCTCCCGAGGGGCGGTGTGGCGAGCGCGCGGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
Db 533 AAGGCTATGGACACCGCTCGCGGTCCCGCGTGTCTCCGGCTACGTGTATCCGCTTGC 592
QY 180 aAspGluGlyMetAspCysAlaLeuAlaValTyrGlyGlyGlyGlyValle 200
Db 593 CGACGAGGGGATGAGTCTGCGGCTCCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTT 652
QY 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 653 GTCACTGTCTCGCGAGCTCAAGGCGTGTTCGAGGAGCAGCTCTGTCTCACTTCTACTC 712
QY 220 rIlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyAspGlyGlyGly 240
Db 713 GGGCTCGATTCTTCTGGGCTATGATCTGCTGCAGTCCGACGAGGCGGAGGTGGGGTGG 772
QY 240 vValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 773 GGTAAACAGTGAAGCTGGTGACTTTCCTCCATGTGGCCGAGGGGTGATGGGGTGAATGACCA 832
QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 943
```

```
Db 833 CAACCTTCCTGGCGGGCTCTGCTAGCTGATCAAGTTGTTTTCTGACATTGTTCCAGAGAC 892
QY 280 rProHisThrGlnProLeuGlyProSer 289
Db 893 TCCTCAGACGCGACCTTTGGGTCTTCT 920

RESULT 5
US-10-042-894A-1
; Sequence 1, Application US/10042894A
; Publication NO. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(806)
US-10-042-894A-1

Alignment Scores:
Pred. No.: 2,13e-158 Length: 1169
Score: 1406.50 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 92.47% Indels: 2
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-1 (1-1169)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 84 ATGCCCGACCTTCACCGCGGAGCACCAGTCGCGGTCACCGCGCTCCGCCAGCAAG 143
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 144 CTGGGCCCCGCTCATCGACGGCTCCGGCTCTTCTACAAGCCGCTCCAGCGCGGACCGT 203
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 204 GGGGAGCAGAGTTCGCTTCTATGAGGCGTTCCTCCGCCAGCGCGCTCCGCCCGCC 263
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 264 ATCCGAGACACCTTCTCCCGCGGTTCACGCGCAGCTCTCTCCCGCGGCGCAG 323
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 324 CCGGGGAGCGCATCCGACCTCTCTCGACGACTCTCTCGCGGGGTTGAGGGGCC 383
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 384 TGCGTCGACAGCATCAAGATCGGCGCATCATCGTGCCACCGAGTTCGCGGAGCCCTAC 443
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 444 ATCCCAAGTACCTCGCCAGGACCGCGGAGCCAGAGCGTTCCTCGGATTCGCGGTC 503
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
```

```
Db 504 TTGC--GTCCAGATCGTCCGCCCGAGGGCCCGTGTGGCGGACGGAGCCCGGAGGTG 561
Qy 161 LyAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 562 AAGGCTATGGACACCGTCGCGCGTCCGCCGCGTGTCTCCGGCGCTACGTTGTCATCCGCTGC 621
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLe 200
Db 622 CGACGAGGGGATGACATGCGCGCTCGCGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTT 681
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220
Db 682 GTACAGCTCGCGGAGCTCAAGGATGTTGGAGGACGACATCTGTTCCACTTCTACTC 741
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyAspGlyGlyG 240
Db 742 GGGCTCGATTCTTCTGGGCTATGATGCTGTCAGTCGCAGCGAGCGGAGGTGGGGTGG 801
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 802 GGTAAACAGTAGTGGTGACTTTTCCCATGTGGCGGAGGTGATGGGGTGGATTGACCA 861
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 862 CNACTTCTGGCGAGCTCTGCTAGCTGATCAAGTTCGTTTCTGACATTGTTCCAGAGAC 921
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 922 TCCTTAGACGAGCCTTTGGGTCTCTTCT 949
```

RESULT 6

```
US-10-042-894A-20
; Sequence 20, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3416
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)...(407)
US-10-042-894A-20
```

```
Alignment Scores:
Pred. No.: 7,43e-146 Length: 3416
Score: 1307.00 Matches: 260
Percent Similarity: 90.34% Conservative: 2
Best Local Similarity: 89.66% Mismatches: 9
Query Match: 85.93% Indels: 20
DB: 14 Gaps: 1
```

US-10-042-894A-8 (1-289) x US-10-042-894A-20 (1-3416)

```
Qy 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 72 ATGCCGACCTCCACCGCGGAGCACCAAGTCCCGGTCTACCGCGCTCCGCGCAGCAAG 131
```

```
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 132 CTGGGCCCCACTCATCGACGACTCTGGGCTCTTCTACAAGCGGCTCCAGGCGCGGACCGT 191
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 192 GGGGAGCAGGAGTCCGCTTCTATGAGGGGTCTTCGCCGCCACCGCCGCTCCGGGCGCG 251
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 252 ATCCGAGACACCTTCTTCCCCCGGTTCCAGCGACCGGACTCTCTCCACCGAGGCGCAG 311
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db 312 CCCGGGAGCGCATCGCACCTCTCTCGACGACTCTCTCGCGGGGTTTGAGGCGCCC 371
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 372 TGGCTCGCAGACATCAAGATCGGTGCATCACGTG----- 406
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 407 -----ACCACGAGCGTTCTGCTCGGATTCGCGGTC 436
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 437 TCCGGCGTCCGAGTCTGTCGCCCGGAGGCGCGTGTGGCGGACGAGCGCCGAGGTG 496
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 497 AAGGCTATGACATTGTCGGCGTCCGCCGCTCTCCGGCGCTACGTTGTCATCGGCTGC 556
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200
Db 557 CGACGAGGGGATGACTGCGCGCTCGCGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTT 616
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSe 220
Db 617 GTACAGCTGCGCGAGCTCAAGCGGTGTTTCGAGGGGCGACACTCTGTTCCACTTCTACTC 676
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyG 240
Db 677 GGGCTCGATTCTTCTGGGCTATGATGCTGTCAGTGCACGAGCGGAGGTGGGGGTGG 736
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 737 GGTAAACAGTAGTGGTGACTTTTGGCCATGTGGCGGAGGTGATGGGGTGGATTGACCA 796
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 797 CNACTTCTGGGCGGCGCTCTGCTAGCTGATCAAGTTGTTTCTGACATTGTTCCAGAGAC 856
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 857 TCCTCAGACGAGCCTTTGGGTCTCTTCT 884
```

RESULT 7

```
US-10-042-894A-15
; Sequence 15, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
```

```
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15
/ LENGTH: 899
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (89)...(424)
US-10-042-894A-15

Alignment Scores:
Pred. No.: 9-8e-138 Length: 899
Score: 1233.00 Matches: 248
Percent Similarity: 88.42% Conservativeness: 4
Best Local Similarity: 87.02% Mismatches: 14
Query Match: 81.07% Indels: 20
DB: 14 Gaps: 1

US-10-042-894A-8 (1-289) x US-10-042-894A-15 (1-899)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB 89 ATGCCCGACCTCCACCGCGGAGCACCAGTCGCGGTCCACCGCGCTCCGCGCAGCAAG 148
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 149 CTGGGCGCCACTATCAGCGCTCTGCGCTCTTCTACAGCGCTCCAGCGCGCGACCGT 208
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 209 GGGGAGCAGAGGTGCGCTTCTATGAGCGTTCCTCGCCACGCGCGCTCCGCGCGCG 268
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 269 ATCCGAGACACCTTCTCCCGCGTTCACGCGCAGCTCTCTCCACCGAGCGCGCAG 328
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 329 CCCGGGAGCGGATCCGATCTCTGTCGCGACCTCTCGCGGGTTTGAAGGCGCC 388
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB 389 TCGTTCGCGACATCAAGATCGGTGCTCATCGTGACCATG-AGCGAT----- 435
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 436 -----CTGCTCGGATTCCACGTC 453
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 454 TCGCGGTCCGAGTCTGCGGCCCGGAGGCGCGTGTGCGGACGAGCGCGCTTAGGGT 513
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB 514 AAGGCTATGGACATTCGGCGTCCGCGGTGCTCCGCGGTGTCATGTCATCCGCTTC 573
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyVal 200
DB 574 CGCGGAGGGGATGAGTTCGCGCTCCGCGCGCGGTGTACGAGGAGAAAGTGGAGTCTT 633
QY 200 uSerGlnLeuArgGluLeuValaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
DB 634 GTCACAGCTGCGGAGCTCAAGCGGTGTTCAGGGGCGAGTCTCTTTCATCTTACTTC 693
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyAspGlyGlyG 240
DB 694 GCGTTCGATTCTTCTGGGTATGATCTCTGTCAGTCGAGCGAGCGGAGGTTGGGTG 753
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAsp 260
DB 754 GGTAAAGTGAAGCTGGTGACCTTCCCATGTGGCCGAGGGGTGATGGGGTGATTGACCA 813
```

```
QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
DB 814 CAACCTTCCTGGCGGCTCTGTAGCTGATCAAGTTTGTTTCTGACATTTGTTCCAGAGAC 873
QY 280 rProHisThrGln 284
DB 874 TCCTTAGACGCGAG 886

RESULT 8
US-10-437-963-4650
/ Sequence 4650, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 4650
/ LENGTH: 1108
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_11514C.1
US-10-437-963-4650

Alignment Scores:
Pred. No.: 1-76e-127 Length: 1108
Score: 1148.50 Matches: 218
Percent Similarity: 84.45% Conservativeness: 21
Best Local Similarity: 77.03% Mismatches: 39
Query Match: 75.51% Indels: 5
DB: 19 Gaps: 3

US-10-042-894A-8 (1-289) x US-10-437-963-4650 (1-1108)

QY 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLysLeu 21
DB 7 TCCGACCTCGCGCCCGCGAGCACCAGGTGGCGGGGCACCGCGCTCCGCCGACCAAGCTG 66
QY 22 GlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGly 41
DB 67 GCGCGCGCTCTCTCGAGCGGCGGGGTCTTCTACAAGCCCTCCAGCGCGGAGCGCGG 126
QY 42 GluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIle 61
DB 127 GAGCAGAGGCGCGCTTCTACCGCGGTTCACCGCGCACCGCGCGTCCGCCCGCGGTC 186
QY 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81
DB 187 CGGGCGCGCTTCTTCCCGCTTCCACGCGCACCCGCTTCTCCCGCGCCCGACGAGCC 246
QY 82 ---GlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB 247 GCGCGCGCGCTTACCGGACATCTCTCGAGACCTCTCTCGCGGCGCTCCCGTCCCC 306
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB 307 TCGTTCGCGAGGTCAAGATCGCGCTCGACGTGCGCGCGCGATCCCGGACCCCTAC 366
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 367 GTCGCAAGTGTCTCGCAAGGACCGAGACACCAGCGCGCTCTCTCGGCTTCGCGCTC 426
```



```

Qy 141 SerGlyValArgValValGlyProGlu---GlyAlaValTrrArgThrGluArgProGlu 159
Db 427 TCCGGCGTCCGGGTGGTTCGATGCCCGGGGGCGCGCGTGTGGCGCCGGACCGGTCCGAG 486
Qy 160 ValLysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerVal 179
Db 487 CTGAAGGGGATCGACGCCCGCGGGGTCCGCGCGGTGTCTCCGCGGCTACGTGTCCACGGC 546
Qy 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyLysGlyGlyVal 199
Db 547 GCGCGGACCGGCTGGAGCTGGCGGCTCGCCGCGCGGTGTACGAGGGGGCGGGCGTC 606
Qy 200 LeuSerGlnLeuArgGluLeuLysAlaTrrPheGluGluGlnThrLeuPheHisPheTyr 219
Db 607 CTGGCTCAGTTCGGGAGCTCAAGGCGTGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
Qy 220 SerAlaSerLeuLeuGlyTyrAsp-----AlaAlaValAlaAlaGlyGly 236
Db 667 TCGCGGTTCGATTCGTCGGCTACGACGCCAATCGCGCGCGCGGTGTCTCCCGGAGGT 726
Qy 237 AspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyGlyGly 256
Db 727 GGAAGCGGCGGTGAAGGCTGAAGCTGTGGAGCTTCGCGCATGTTCGACGATGGGAGCGG 786
Qy 257 ValLeuAspHisAsnPheLeuGlyGlyLeuCysSerLeuLeuLysPheValSerAspIle 276
Db 787 GTGATTCACCACTTCTTGGCGGGGTCTGTCTCGCTCATCAAGTTTCATCGCGGACATT 846
Qy 277 ValProGlu 279
Db 847 GTCGCGAG 855

```

# RESULT 9

```

US-10-437-963-4649
; Sequence 4649, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 4649
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11513C.1
US-10-437-963-4649

```

Alignment Scores:

| Pred. No.:             | Length:  | Matches:         |
|------------------------|----------|------------------|
| Score:                 | 1.22e-91 | 696              |
| Percent Similarity:    | 847.50   | 173              |
| Best Local Similarity: | 66.31%   | Conservative: 14 |
| Query Match:           | 61.35%   | Mismatches: 28   |
|                        | 55.72%   | Indels: 67       |
|                        | 19       | Gaps: 3          |

US-10-042-894A-8 (1-289) x US-10-437-963-4649 (1-696)

```

Qy 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLeu 21
Db 7 TCCGACCTGCCCGCGCGGACCAAGGTGGCGGGGACCGCGGTCCGCGGACCAAGCTG 66

```

```

Qy 22 GlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGly 41
Db 67 GGGCCGCTCGTCGACGCGGAGGGGCTCTTCTACAAGCCCTCCAGGCGGGGAGCGGG 126
Qy 42 GluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIle 61
Db 127 GAGCAGGAGCGCGCTTTTACGCGCGGTTCACCGCGCACCCGCGCTCCCGCCCGG--- 183
Qy 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81
Db 183 ----- 183
Qy 82 GlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaProCys 101
Db 183 ----- 183
Qy 102 ValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIle 121
Db 183 ----- 183
Qy 122 AlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSer 141
Db 184 -----TGCTTCGCCAAGGACCCGAGACCCAGCGCGCTCTCGGCTTCGCGCTCC 237
Qy 142 GlyValArgValValGlyProGlu---GlyAlaValTrrArgThrGluArgProGluVal 160
Db 238 GGGCTCGGGTGGTTCGATGCCCGGGGCGCGCGTGTGGCGCCGAGCCGCTCGGAGCTG 297
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
Db 298 AAGGATCGACCGCGCGGGTCCGCGCGTCTCCGCGCTACGTTGTCACGCGGCGGC 357
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrrGlyGlyLysGlyGlyValLeu 200
Db 358 GCGACGCGCTGGACTGCGCGCTCGCGCGCGGTGTACGAGGGGAGGGGCGGCTCTG 417
Qy 201 SerGlnLeuArgGluLeuLysAlaTrrPheGluGluGlnThrLeuPheHisPheTyrSer 220
Db 418 GCTCAGCTCGGGGAGCTCAAGGCGTGTTCGAGGAGCAACCCCTGTACCACTTCTACTCG 477
Qy 221 AlaSerIleLeuLeuGlyTyrAsp-----AlaAlaValAlaAlaGlyGlyAsp 237
Db 478 GCGTCGATTCGTCGCTACGACGCAATGCGCGCGCGCGCTGTCTCCCGAGGTGGA 537
Qy 238 GlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 538 AGCGCGGGTGAAGGTGAAGCTTGGTGGACTTCGCGCATGTCCAGCATGGGAGCGGGTG 597
Qy 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 598 ATTGACCACTTCTTGGCGGGCTCTGCTCGCTCATCAAGTTTCATCGCGGACATTGTC 657
Qy 278 ProGlu 279
Db 658 GCGGAG 663

```

## RESULT 10

```

US-10-042-894A-17
; Sequence 17, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12

```

```
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(643)
; OTHER INFORMATION: n = A, T, C or G
US-10-042-894A-17

Alignment Scores:
Pred. No.: 5,26e-87 Length: 643
Score: 808.50 Matches: 162
Percent Similarity: 93.14% Conservative: 1
Best Local Similarity: 92.57% Mismatches: 10
Query Match: 53.16% Indels: 3
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-042-894A-17 (1-643)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 117 ATCCCGACCTCCACCGCGGAGCACCAGTCGCGGTCCCGCGCTCCGCGCAAG 176
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 177 CTGGGCGCGTCTATCGACGCGTCCGCGCTCTTCTACAAGCGGTCCAGCGCGCGACCGT 236
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 237 GGGGACACGAGGTGCGCTCTCTATGAGCGTCTCCGCCACCGCGCGTCCGCGCGCG 296
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 297 ATCCGAGACACCTTCTTCCCGCGTTCACGCGACGCGACTCTCCGCCACCGCGCGCG 356
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db 357 CCGGGGAGCGGATCCGACCTCTGCTCGACGACCTCTCTCGCGGGTTTGAGGGCGCC 416
QY 101 CysValAlaAspIleValIleGlyAlaIleThrProProSerSerProGluProTyr 120
Db 417 TGGTTCGACACATCAAGATCGCGGCATCATCGTGGCCACCGAGTTCGCGGAGCGCTAC 476
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 477 ATCGNCAAGTACCTNGCCAAAGGACCGCGGACCACGAGCGTTCTGCTCGGATTCGCGTC 536
QY 141 SerGlyValArgValGlyProGluGlyAlaValThrArgThrGluArg-ProGluVa 160
Db 537 TTGC--GTCCGAGTCTGCGGCCCGAGGGCGCGTGTGGCGGACGAGCGCGCGGGGT 594
QY 160 lLysAlaMetAspThr-AlaGlyValArgValLeuArg 173
Db 595 GAANGCTATGACACCCCGTCGGNGNCCGGCGGTGCTTCGG 635

RESULT 11
US-10-042-894A-13
; Sequence 13, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE OF INVENTION: 1286
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
```

```
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Parthenium argentatum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(908)
US-10-042-894A-13

Alignment Scores:
Pred. No.: 7,93e-79 Length: 1020
Score: 742.50 Matches: 149
Percent Similarity: 67.83% Conservative: 45
Best Local Similarity: 52.10% Mismatches: 79
Query Match: 48.82% Indels: 13
DB: 14 Gaps: 5

US-10-042-894A-8 (1-289) x US-10-042-894A-13 (1-1020)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 24 CTCAAGGCCCCAGATCATCAGGTTCGTGACATGAAGCTGGGCTCGGGAAGCTTGGCCCA 83
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 84 CTCATTGATGATTCAGGCGCGTTCACAAACCACTGCAGGGTGATAACCGTGGTCAGAA 143
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 144 GAAGTAGCCCTTTATGAATCATTTCTTCTACAATAATATTCAGAACACACATACGC--- 200
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 201 AAATCTCTTCTATATATTATGACCAACAAATCATG-----AAGGCATCCACTGGCTCT 254
QY 84 ProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaProCysValAla 103
Db 255 GACCATCTCTACATGGTGTGCAAGATCTTACATCAGCTCATGTCAACCCATCTGTAAATG 314
QY 104 AspIleLysIleGlyAlaIleThrProProSerSerProGluProTyrIleAlaLys 123
Db 315 GACATCAAAATCGGTCCAGAACATGGCGCCGACGAAGCTTCCGAGCGGTACATTGCAAA 374
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 375 TGCCTAAAAAGGATAGGGAAGACACAGTATTCCATTGGGATTCAGGATCTCCGGGCTG 434
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
Db 435 CAAGTCTATATCGATGGGTGAGGCTTTTATAAGCCTCATAGAAATTTACATGGCTAA 494
QY 163 MetAspThrAlaGlyValArgValLeuArgValGlyTyrValSerSerValAlaAspGlu 182
Db 495 ACCGGCCCGAGTGTGTAGACTACTCTTAGAAATTTGTTCTTAACCCGCTCGCA 554
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
Db 555 GAGATGGAATCGGCACAGGCCCTAGCCCGGATTTCTTTAGCATCTTTTGTATTGTGT 614
QY 195 GlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThr 214
Db 615 GGGCCTAATGGGATATTAGCTCACTGATGGAATTTGAAGACATGGTTTGAAGATCAAA 674
QY 215 LeuPheHisPheSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAla 234
Db 675 ATTTACCACTTCTATGCTTGTCTTTTGTTCATCTTTTGAAGAGGTGGTGTAAAA 734
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGly 254
```

```
Db 735 GGTGCT---CGGTCAACAGCAGAGTCAAACTTATTGATTGCTCATGTTACAGATGGT 791
Qy 255 AspGlyValIleAspHisAsnPheLeuGlyGlyLeuCySerLeuIleLysPheValSer 274
Db 792 AATGGGTGTTATTGATCAAAATTTCTGGGTGGGCTCTGTTCTTTGATAAAGTTCAATTTCT 851
Qy 275 AspIleValProGluThr 280
Db 852 GACATACTTTCGGAGACA 869

RESULT 12
US-10-042-894A-11
; Sequence 11, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1195
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)...(1048)
US-10-042-894A-11

Alignment Scores:
Pred. No.: 2.44e-76 Length: 1195
Score: 722.50 Matches: 149
Percent Similarity: 66.31% Conservative: 38
Best Local Similarity: 52.84% Mismatches: 84
Query Match: 47.50% Indels: 11
DB: 14 Gaps: 6

US-10-042-894A-8 (1-289) x US-10-042-894A-11 (1-1195)
Qy 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 119 CTCAAGTCCGGATCATCAAGTCGCCGGTCCAGCGGGGAGACGGGGGAAAGCTGGGGCCA 178
Qy 24 LeuIleAspGlySerGlyLeuPheTyLysProLeuGlnAlaGlyAspA:gglyGluHis 43
Db 179 CTGTGGATGATTCGGGCCCGCTTCTAAGCCTCTCCAGAGCGATCATCGCGGAGACAG 238
Qy 44 GluValAlaPheTyGluAlaPheSerAlaHisAlaValProAlaAargIleArgAsp 63
Db 239 GAAGTGGCCCTTTACGAGTCAATCTATTTCATACCGAGATCCCGAGTCATTCGC--- 295
Qy 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 296 AATTTCTTTCTCGGTTTACGGAACTAAGACTATT-----GAGGGCTCTGATGGATCG 349
Qy 84 ---ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCyVal 102
Db 350 GGTCCTCAACCTCACTGGTTCTGGAGGATCTGCTCGGGTGCACGAAACCCATCTCTC 409
Qy 103 AlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyIleAla 122
Db 410 ATGGACATCAAGACTGGATCCAGAACATGGTATCCGGAGGCGCTCTGAGGATACATCAA 469
Qy 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
```

```
Db 470 AAGTCTTAGAGAAAGATGAAATAGCAACAGCGTTTCATTGGGTTTATAGGATTTCTGGG 529
Qy 143 ValArgValVal---GlyProGluGlyAlaValATpArgThrGluArgProGluValLys 161
Db 530 CTAAAGGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAGAAGTTGTTTAT 589
Qy 162 AlaMetAspThrAlaGlyValArgValLeuArgArgTyValSerSerValAlaAsp 181
Db 590 AGCTTAAATCGGACCGGTGTCAAGTCGGCTCTGAGGAAGTTTGTCTTCTTCCAACTGTCT 649
Qy 182 GluGly-----MetAspCysAlaLeuAlaValTyArgTyGlyGlyLysGly 197
Db 650 CTGGGTCCAAATGTGGATCCGATTTGTTGATGTCATCAAAAGTTTACTGTCTCCCGGGT 709
Qy 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTppPheGluGluGlnThrLeuPheHis 217
Db 710 GGAATTTGGCAAAATGCTTCAGCTGAAGGAATGGTTTGAGGTTTCAGAGCAATATAC 769
Qy 218 PheTySerAlaSerIleLeuLeuGlyTyArgAlaAlaAlaValAlaAlaGlyAsp 237
Db 770 TTCTATTCTGTCTCACTCATTTATATGACAGGAGTCTGCTTTGGACGGC----- 823
Qy 238 GlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 824 TGTGCACACCCGAAAGTTAAACTGGTGGACTTTGCACATGTGATGGATGCCACGGCGTG 883
Qy 258 IleAspHisAsnPheLeuGlyGlyLeuCySerSerLeuIleLysPheValSerAspIleVal 277
Db 884 ATCGATCAACAACCTTCTTGGGTGGCCTCTGTCTGTAAATCAAGTTTATACGTGACATTGCT 943
Qy 278 ProGlu 279
Db 944 GATGAA 949

RESULT 13
US-10-042-894A-9
; Sequence 9, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafaleki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)...(851)
US-10-042-894A-9

Alignment Scores:
Pred. No.: 5.01e-67 Length: 1105
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservative: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 14 Gaps: 8

US-10-042-894A-8 (1-289) x US-10-042-894A-9 (1-1105)
```

```
QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAsp 26
Db 24 CCGGAGCACCAAGTGGCCGGGCAAGGCAAGGACGGAATCTGGGCCCACTCGTCGAC 83
QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42
Db 84 GATTTTGGAAAAATTCACAAAGCCCTCCAGACCAAAAGACGACGACACCGCGGCTCC 143
QY 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArg 62
Db 144 ACCGAATCTCTCTTTACACCTCTCGCGGCC--GCCGCCACGACTACTCCATCCGC 200
QY 63 AspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly 82
Db 201 ---TCCTTCTTCCCGCCCTTTCAGGACCCGCTCTCGACGCTCCGACGCGCTCCGCT 257
QY 83 GluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db 258 ---CCGACACCTCACCTGGTCTCGAGGACCTCTCTGGGGTACTCTCAAAACCCCTCGTC 314
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db 315 ATGGAGCTAAAGATCGGCTCCAGACCTGGACCTGGGAGACTCCGAGGACTACATCTGC 374
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 375 AAGTGGCTGAAGAGGACAGAGAGTCTCTAGCTTGCCTTGGTTTCAGAAATCTCGGA 434
QY 143 ValArgValValGlyProGluGlyAlaValTTPArgThrGluArgProGluValLysAla 162
Db 435 GTCAAG-----GACTCTATCTCTCTCTGGGAACCTACCAGGAAATCTCTCCAGTGT 485
QY 163 MetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer----- 178
Db 486 CTATCGCCCATGGTGTGACCTGTCTTCAACAGTTCTGTTTCTCTTAATATCAAC 545
QY 179 ValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 198
Db 546 CATGATGATCATCATCCGATTCGGTTTCGCACACGAGGAGTCTAC-----GCCGCC 596
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGlnThrLeuPheHisPhe 218
Db 597 GTTTTGGAGCGCTTGCAAGAGCTCAAGGACTGGTTTCGAGGTTTCAGACGCGTATCACTTC 656
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGly 238
Db 657 TATTTCTGTCTGTCTTGTGTGTACGAGAGGAT-----CTAGGGAAGGGAAGCT 710
QY 239 GlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIle 258
Db 711 ACCAACCCCTCTGTCAAACTCGTTGACTTTGCACACGTTGTGAGACGGAACGGTGTCA 770
QY 259 AspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 771 GATCACAACTCTTGGGTGGCTTCTTCTTCATCAAGTTCTCAAGGATATCTTA 827
```

## RESULT 14

```
US-10-424-599-106467
; Sequence 106467, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 106467
; LENGTH: 1345
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67156C.1
; US-10-424-599-106467
```

```
Alignment Scores:
Pred. No.: 6,478-67 Length: 1345
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservatives: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 18 Gaps: 8
```

US-10-042-894A-8 (1-289) x US-10-424-599-106467 (1-1345)

```
QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAsp 26
Db 229 CCGGAGCACCAAGTGGCCGGGCAAGGCAAGGACGGAATCTGGGCCCACTCGTCGAC 288
QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42
Db 289 GATTTTGGAAAAATTCACAAAGCCCTCCAGACCAAAAGACGACGACACCGCGGCTCC 348
QY 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArg 62
Db 349 ACCGAACTCTCTTTTACACCTCTCTCGCGGCC--GCCGCCACGACTACTCCATCCGC 405
QY 63 AspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly 82
Db 406 ---TCCTTCTTCCCGCCCTTTCAGGACCCGCTCTCGACGCTCCGAGGAGTCTCCGCT 462
QY 83 GluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db 463 ---CCGACACCTCACCTGTCTCGAGGACCTCTCTCGCGTCTCCAAACCCCTCCGTC 519
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db 520 ATGGAGCTAAAGATCGGCTCCAGACCTGGACCTGGGAGACTCCGAGGACTACATCTGC 579
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 580 AAGTGGCTGAAGAGGACAGAGAGTCTCTAGCTTGCCTTGGTTTCAGAAATCTCGGA 639
QY 143 ValArgValValGlyProGluGlyAlaValTTPArgThrGluArgProGluValLysAla 162
Db 640 GTCAAG-----GACTCTATCTCTCTCGGGAACCTACCAGGAAATCTCTCCAGTGT 690
QY 163 MetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer----- 178
Db 691 CTATCGCCCATGGTGTGACCTGTCTTCAACAGTTCTGTTTCTCTTAATATCAAC 750
QY 179 ValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 198
Db 751 CATGATGATCATCATCCGATTCGCTTTCGCAACGAGGAGTCTAC-----GCCGCC 801
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGlnThrLeuPheHisPhe 218
Db 802 GTTTTGGAGCGCTTGCAAGAGCTCAAGGACTGGTTTCGAGGTTTCAGACGCGTATCACTTC 861
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGly 238
Db 862 TATTTCTGTCTGTCTTGTGTGTACGAGAGGAT-----CTAGGGAAGGGAAGCT 915
QY 239 GlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIle 258
Db 916 ACCAACCCCTCTGTCAAACTCGTTGACTTTGCACACGTTGGTGGACGGAACGGTGTCA 975
QY 259 AspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 976 GATCACAACTCTTGGGTGGCTTGTCTTCTTCATCAAGTTCTCTCAAGGATATCTTA 1032
RESULT 15
```

US-10-425-115-74436  
; Sequence 74436, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 74436  
; LENGTH: 776  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(776)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_167884C.1  
US-10-425-115-74436

Alignment Scores:  
Pred. No.: 5,73e-63 Length: 776  
Score: 609.00 Matches: 117  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.15% Mismatches: 0  
Query Match: 40.04% Indels: 0  
DB: 20 Gaps: 0

US-10-042-894A-8 (1-289) x US-10-425-115-74436 (1-776)

|    |     |  |     |
|----|-----|--|-----|
| QY | 172 | LeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla    | 191 |
| DB | 76  | GTCCGGCGCTACGTGTCNTCCGTTGCCGACGAGGGGATGGACTGTGGCTCGCGGGCG    | 135 |
| QY | 192 | ValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaIrrPheGlu    | 211 |
| DB | 136 | GTGTACGGAGGAAAGGTGGAGTCTTGTACAGCTGCCGAGCTCAAGGCGTGGTTCGAG    | 195 |
| QY | 212 | GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAla    | 231 |
| DB | 196 | GACGAGACTCTGTCCACTTCTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGTCA    | 255 |
| QY | 232 | ValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisVal    | 251 |
| DB | 256 | GTCCGACGAGCGGAGATGGGGGTGGGTGACCGGTGAGCTTGTGGACTTTGCCCATGTG   | 315 |
| QY | 252 | AlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLys | 271 |
| DB | 316 | GCCGAGGGTGATGGGGTGATTGACCACAACCTTCTGGCGGGCTCTGCTCGCTGATCAAG  | 375 |
| QY | 272 | PheValSerAspIleValProGluThrProHisThrGlnProLeuGlyProSer       | 289 |
| DB | 376 | TTGCTTCTGACATTGTTCCGAGACTCTCTCATACGAGCCTTTGGGTCTCTTCT        | 429 |

Search completed: June 17, 2005, 09:05:18  
Job time : 2137 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 03:46:13 ; Search time 205 Seconds

(without alignments)  
2306.751 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVACHRASASK.....IKFVSDIVPTHTQPLGPS 289

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111951\_8314/app\_query.fasta\_1.455  
-DB=Issued Patents NA -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894 @CGN\_1\_1\_69 @runat\_15062005\_111951\_8314 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 193   | 12.7        | 1341   | 4  | US-09-023-655-809    |
| 2          | 189   | 12.4        | 1748   | 4  | US-09-620-312D-712   |
| 3          | 182   | 12.0        | 1863   | 4  | US-09-774-528-103    |
| 4          | 178.5 | 11.7        | 1893   | 4  | US-09-774-528-102    |
| 5          | 145.5 | 9.6         | 1781   | 4  | US-09-949-016-4634   |
| 6          | 139   | 9.1         | 963    | 4  | US-09-248-796A-5436  |
| 7          | 113.5 | 7.5         | 1899   | 4  | US-09-252-991A-6525  |
| 8          | 112.5 | 7.4         | 4496   | 4  | US-09-949-016-2348   |
| 9          | 112.5 | 7.4         | 4505   | 4  | US-09-949-016-323    |
| 10         | 108   | 7.1         | 14555  | 4  | US-09-902-540-1096   |
| 11         | 107.5 | 7.1         | 954    | 4  | US-09-248-796A-4664  |
| 12         | 107.5 | 7.1         | 984    | 4  | US-09-252-991A-12872 |

|      |       |     |         |   |                      |                   |
|------|-------|-----|---------|---|----------------------|-------------------|
| C 13 | 107.5 | 7.1 | 4078    | 4 | US-09-016-434-1132   | Sequence 1132, Ap |
| C 14 | 106.5 | 7.0 | 6000    | 1 | US-08-348-006B-6     | Sequence 6, Appli |
| C 15 | 106.5 | 7.0 | 6000    | 2 | US-08-800-825A-6     | Sequence 6, Appli |
| C 16 | 106.5 | 7.0 | 6000    | 3 | US-09-158-657-6      | Sequence 6, Appli |
| C 17 | 106.5 | 7.0 | 6000    | 5 | PCT-US94-10166-6     | Sequence 6, Appli |
| C 18 | 106   | 7.0 | 4411529 | 3 | US-09-103-840A-1     | Sequence 1, Appli |
| C 19 | 105.5 | 6.9 | 36412   | 4 | US-08-311-731A-132   | Sequence 132, App |
| C 20 | 104   | 6.8 | 1917    | 2 | US-08-637-899-2      | Sequence 2, Appli |
| C 21 | 103.5 | 6.8 | 1896    | 4 | US-09-252-991A-11331 | Sequence 11331, A |
| C 22 | 101.5 | 6.7 | 8332    | 4 | US-09-902-540-927    | Sequence 927, App |
| C 23 | 101   | 6.6 | 1161    | 4 | US-09-902-540-7636   | Sequence 7636, Ap |
| C 24 | 101   | 6.6 | 2460    | 4 | US-09-902-540-4178   | Sequence 4178, Ap |
| C 25 | 101   | 6.6 | 12419   | 4 | US-09-902-540-1059   | Sequence 1059, Ap |
| C 26 | 101   | 6.6 | 22301   | 4 | US-09-902-540-1208   | Sequence 1208, Ap |
| C 27 | 100.5 | 6.6 | 960     | 4 | US-09-252-991A-12565 | Sequence 12565, A |
| C 28 | 100.5 | 6.6 | 2314    | 3 | US-09-144-914-3      | Sequence 3, Appli |
| C 29 | 100.5 | 6.6 | 2568    | 4 | US-09-949-016-1555   | Sequence 1555, Ap |
| C 30 | 100.5 | 6.6 | 3446    | 4 | US-09-620-312D-653   | Sequence 653, App |
| C 31 | 100.5 | 6.6 | 17173   | 4 | US-09-902-540-1122   | Sequence 1122, Ap |
| C 32 | 100.5 | 6.6 | 43507   | 4 | US-09-949-016-13297  | Sequence 13297, A |
| C 33 | 100   | 6.6 | 1425    | 4 | US-09-489-039A-1509  | Sequence 1509, Ap |
| C 34 | 100   | 6.6 | 1491    | 4 | US-09-252-991A-6021  | Sequence 6021, Ap |
| C 35 | 100   | 6.6 | 2016    | 4 | US-09-489-039A-1492  | Sequence 1492, Ap |
| C 36 | 99.5  | 6.6 | 5663    | 4 | US-09-902-540-839    | Sequence 839, App |
| C 37 | 99.5  | 6.5 | 1681    | 4 | US-09-270-767-1662   | Sequence 1662, Ap |
| C 38 | 99.5  | 6.5 | 1681    | 4 | US-09-270-767-16944  | Sequence 16944, A |
| C 39 | 99    | 6.5 | 1470    | 4 | US-09-489-039A-6625  | Sequence 6625, Ap |
| C 40 | 99    | 6.5 | 19068   | 4 | US-09-902-540-1123   | Sequence 1123, Ap |
| C 41 | 99    | 6.5 | 23856   | 4 | US-09-949-016-15732  | Sequence 15732, A |
| C 42 | 98.5  | 6.5 | 7201    | 4 | US-09-902-540-914    | Sequence 914, App |
| C 43 | 98    | 6.4 | 12194   | 4 | US-09-902-540-1091   | Sequence 1091, Ap |
| C 44 | 98    | 6.4 | 23677   | 4 | US-09-902-540-1218   | Sequence 1218, Ap |
| C 45 | 98    | 6.4 | 4403765 | 3 | US-09-103-840A-2     | Sequence 2, Appli |

#### ALIGNMENTS

#### RESULT 1

US-09-023-655-809  
; Sequence 809, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:



```
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 809:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT01
; CLONE: 753522
US-09-023-655-809

Alignment Scores:
Pred. No.: 2,37e-12 Length: 1341
Score: 193.00 Matches: 88
Percent Similarity: 40.36% Conservative: 48
Best Local Similarity: 26.11% Mismatches: 110
Query Match: 12.69% Indels: 94
DB: 4 Gaps: 10

US-10-042-894A-8 (1-289) x US-09-023-655-809 (1-1341)
QY 8 GluHisGlnValAlaGlyHisArgAlaSerAlaSerLys-LeuGlyProLeuIleAspG1 27
DB 176 AACATCATGCTTAGAAGAACGAAAGAACCCCTAAGGACTGGGTGGTCACACGATTA 235
QY 27 ySerGlyLeuPheTyLysPro----- 34
DB 236 AAGAGGAGAAATGAAGAGCCATAAGTTAGTAAGAACGAATTTGAGTGGCTAAAGAAATCTG 295
QY 35 -----LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyRGlualaph 51
DB 296 AAGCTTTGTACTACACTGTAGAGAAGAGGGAATATAAG-----TT 337
QY 51 eSerAla-----HisAlaAlaValProAlaArgIleArgAspTh 64
DB 338 CCAGCTTAACACTATAACCTTGGAGCATGAATGTCCACGACACAGTTACAGAGAA 397
QY 64 rPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluPr 94
DB 398 T-----GAAGGAGAAATGCAAGAGCATCGGAACC 424
QY 84 o-HisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysValAlaA 104
DB 425 AGTACAAATTTATCTTACTGGAACCTGACTTCCCGCTATGAGTGCCTTGTGTCCTTG 484
QY 104 spileLysIleGlyAlaIleThrTrpProSerSerProGluProTyRlleAlaLysC 124
DB 485 ACCTCAAGATGGGCACACACACATGGTGTATGCTTCAGAGGAGAGGCCAACCC 544
QY 124 ysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgValSerGlyValA 144
DB 545 AGATCCGAAAATGTCAGCAGACACATCTGCAGTCATTGGTGGCTGTGTGGCATGC 604
QY 144 rgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAlaMetA 164
DB 605 AGGTGTAC-----CAAGCAGGCAG-TGGCAGCTCATGTTTCATGAACA-AGTACCATGG 656
QY 164 spThrAlaGlyValArgArgValLeuArgArgTyRValSerSerValAlaAspGluGlyM 184
DB 657 ACGAAGCATTCGGTCAGGGCTTCAGAGGACCTTTCCAGTCTTCCACATGGGGG 716
QY 184 etAspCysAlaLeuAlaAlaValTyRgLyLysGlyValLeuSerGlnLeuA 204
DB 717 GTACCTGGCGCGTG-----AACTCCTG-GGCCCTGTGTCTCAAGAAGCTGA 760
QY 204 rgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyRSerAlaSerIleL 224
DB 761 CTGAGCTCAAGGAGTGTGGAGCGACGAGAGTCTACCGCTTCTACTCAAGCTCCCTGC 820
QY 224 euLeuGlyTyRAsp----- 228
```

```
Db 821 TGGTCATTATGATGGCAAGGAGCGCCCGAGTGGTCTCTGGACTCAGATGCTGAGGATT 880
QY 229 -----AlaAlaAlaValAlaAlaGlyGly----- 236
Db 881 TGGAGGACCTGTGACAGGAATCAGCTGATGAGTCTGCTGGTGCCTATGCTACAAACCCA 940
QY 237 --AspGlyGlyValThrValLysLeuValAspPheAlaHis----- 250
Db 941 TCGGCGCCAGCTCTGTAGATGTGCGCATGATGATGATTTGCACACACACCTCGAGGCTGT 1000
QY 251 -----ValAlaGluGlyAspGlyValIleAspHisAsnPhLeuGlyG 265
Db 1001 ATGGCGAGACACCTGTGGTCATGAGGCCAG-----GATGCTGCTATATCTTCG 1051
QY 265 lyLeuCySerLeuIleLysPheValSerAspIleValProGluThr 280
Db 1052 GGCTCCAGAGCCTGATAGACATTTGTACAGAGATAAGTAGGAGAGT 1098

RESULT 2
US-09-620-312D-712
; Sequence 712, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 712
; LENGTH: 1748
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)..(1467)
US-09-620-312D-712

Alignment Scores:
Pred. No.: 1,04e-11 Length: 1748
Score: 189.00 Matches: 71
Percent Similarity: 45.27% Conservative: 39
Best Local Similarity: 29.22% Mismatches: 83
Query Match: 12.43% Indels: 53
DB: 4 Gaps: 6

US-10-042-894A-8 (1-289) x US-09-620-312D-712 (1-1748)
QY 78 GluAlaGlnProGlyGluPro-HisProHisLeuValLeuAspLeuAlaGlyPh 97
Db 765 GAATCAAAGCATCGGAACCCAGTACAAATTTATCTTACTTGGAAACCTGACTTCCCGCTA 824
```

```
QY 97 eGlnAlaProCysValAlaAspIleValIleGlyAlaIleThrTrpProProSerSerPr 117
Db 825 TGAGGTCCTGTGTCTTCCACCTCAAGATGGGCACACGACACATGATGATGCTTC 884
QY 117 oGluProTyrlleAlaIysCysLeuAlaIysAspArgGlyThrThrSerValLeuLeuG1 137
Db 885 AGAGGAGAGCGACCAACAGATCCGANAATGCACGACGACACATCTCAGTCATGG 944
QY 137 yPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluAr 157
Db 945 TGTGCGTGTGTGGCATGAGGTGTAC-----CAAGCAGGCGAG-TGGCGAGTCATGT 997
QY 157 gProGluValIysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTrpValSe 177
Db 998 TCATGAACA-AGTACCATGACGGAAGCTATCGGTGCGAGGCTTCAGAGGAGGACATTTTC 1056
QY 177 rSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyIysG1 197
Db 1057 CAGTTCTTCACATGGCGGTACTCTGCGCGTG-----AACTCCTG-GG 1100
QY 197 yGlyValLeuSerGlnLeuArgGluLeuIysAlaTrpPheGluGluGlnThrLeuPheHi 217
Db 1101 CCCTGTGCTCAAGAAGCTACTGAGCTCAAGGCAGTGTGTGAGCGACAGGAGTCTTACCG 1160
QY 217 sPheTySerAlaSerIleLeuLeuGlyTyAsp----- 228
Db 1161 CTTCTACTCAAGTCCCTGTGTGTTATTTATGATGGCAAGGCGCGCCGAACTGGTCT 1220
QY 229 ----- 235
Db 1221 GGACTCAGATGCTGAGGATTTGGAGGACCTGTACAGGATCAGCTGATGATGCTGTGG 1280
QY 235 yGly-----AspGlyGlyValThrValIysValIysLeuValAspPheAl 249
Db 1281 TGCCTATGCTTACAAACCCATCGCGCGCAGCTCTGTAGATGTGGCATGATGCACTTGC 1340
QY 249 aHis-----ValaGluGlyAspGlyValI1 258
Db 1341 ACACACACCTGCAGGCTGTATGGCGAGGACACCGTGTGATGAGGGCGCAG----- 1392
QY 258 eAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValPr 278
Db 1393 -GATGTGGTATATCTTCGGGCTCCAGAGCCTGATAGACATGTTGCACAGATAGTGA 1451
QY 278 oGluThr 280
Db 1452 GGAGAGT 1458
```

## RESULT 3

```
US-09-774-528-103
; Sequence 103, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
```

```
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 103
; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (367)..(1593)
US-09-774-528-103

Alignment Scores:
Pred. No.: 7,4e-11 Length: 1863
Score: 182.00 Matches: 67
Percent Similarity: 44.53% Conservative: 47
Best Local Similarity: 26.17% Mismatches: 82
Query Match: 11.97% Indels: 60
DB: 4 Gaps: 10
```

US-10-042-894A-8 (1-289) x US-09-774-528-103 (1-1863)

```
QY 62 ArgAspThrPhePhePro-----ArgPheHisGlyThrArgLeuLeu 75
Db 868 AGGAAGAGCTTCAACCCGTGGGCGCTTGCATATGCCACGAGCCACCTGACCCGCTGTC 927
QY 76 ProThrGluAlaGlnProGlyGluProHisProHisValLeuValLeuAspLeuLeuAla 95
Db 928 -----TCCGAGTACCAGAGAACAAAGCGGTTCTTGTGTGGAAATGATGTGTC 978
QY 96 GlyPheGlnAlaProCysValAlaAlaAspIleLysIleGlyAlaIleThrTrpProSer 115
Db 979 CAGTACACGATCCCTGTGTCTGTGGATCTGAAGATGGGACCCGCGAGCAGCGGATGAT 1038
QY 116 SerProGluProTyrlleAlaIysCysLeuAlaLysAspArgGlyThrThrSerValLeu 135
Db 1039 GCATCGAGGAGAGAGAGGCGCCACATCAGGAAGTGTGCGCAGAGACCTCAGCCTGC 1098
QY 136 LeuGlyPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpArgThr 155
Db 1099 CTGGTGTGGCATCTGCGGCGATGCGAG-----GTTTATCAACA 1137
QY 156 GluArgProGluValIysAlaMetAspThrAlaGlyValArgValLeuArgArgTy 175
Db 1138 GATAAGAAGTACTTCTCTGCAAGAC-----AAGTACTAT 1173
QY 176 ValSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGly 195
Db 1174 GGAAGAAAACTCTCAGTGGAGGGGTTTCAGACAAAGCCCTCTATCAGTTCTACATAATGA 1233
QY 196 Lys-----GlyGlyValLeuSerGlnLeuArgGluLeuLeuAla 208
Db 1234 AGCCACTCGGAGGAGGCTCTCGGAGCCCATCTGCACGAGCTCCGGGCGCTCTCTCT 1293
QY 209 TrpPheGluGluGlnThrLeuPheHisPheTySerAlaSerIleLeuLeuGlyTyAsp 228
Db 1294 GTCATTAGGAGCCAGATTTCATACCGCTTCTATTCAGCTCTCTCTCTGTGTCATGAT 1353
QY 229 -----AlaAlaAlaVal 232
Db 1354 GGGCAGGAACCAACAGAAAGAGCCCGCAGCGCATCTCAGAGGCTCCCGAGGCA 1413
QY 233 AlaAlaGlyGlyAspGlyGlyValThr-----ValIysValAspPheAla 249
Db 1414 GCCCAGGTAGCTCTCCCGGTGTCTCACCAAGGTTGACATCCGCATGATGATTTGCT 1473
QY 250 HisValAla---GluGly-----AspGlyValIleAspHis 260
Db 1474 CATAACCATACAGGCTACTGGAATGACACACACCTACGATGGACCA---GACCT 1530
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276
Db 1531 GGCTATATTTTGGCCTGGAAACCTCATCAGGATCTCTCAGGATATC 1578
```



Db 752 CTCAGGACCTGCTCGATGCTTCGAGGACCTTGTGTGCTGACGTGCAAAATGGGGTC 811  
QY 110 IleThrTrp----- 112  
Db 812 AGGACTTACCTAGAGGAGGAGTGACCAAGGCCGCTGAGCGGCCCAAGCTGCGGAGGAC 871  
QY 113 -----ProPheSerPro-----GluProTyrIleAla 122  
Db 872 ATGTACAAGAAATGCTGGCGGTGGATCTCTGAAGCTCCACGAGGAGGAGGACGGCGCAG 931  
QY 123 LysCysLeuAlaLysAspArgGly-----ThrThrSerValLeu 135  
Db 932 CGCGCGTCAACCAAGCGCGCTACATGCTGCGGCGGAGGAGGATCAGCTCCAGCACCACC 991  
QY 136 LeuGlyPheArgValSerGlyValArgVal-----GlyProGluGlyAlaValTrpArg 154  
Db 992 CTCGGCTTCGCGATCGAGGCGATCAAGAAAGCGGAGCGGCTCTGCGAGCACCAGCTTCNAG 1051  
QY 155 ThrGluArgProGluValLysAlaMetAsp-----ThrAlaGlyValArg 169  
Db 1052 ACTACCGAAGCGGAGGAGGCTGCTCGCGCTCTTTGAAGAGTTTGTGCAAGAGATGAG 1111  
QY 170 ArgValLeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAla 189  
Db 1112 GAAGTCTGAGCGGTAT----- 1129  
QY 190 AlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaTrp 209  
Db 1130 -----CTGAACCGCTGCGAGCAGATCCGGGACACC 1159  
QY 210 PheGluGlnThrLeuPheHisPheTyr-----SerAlaSerIleLeuGly 226  
Db 1160 CTGAGGATCCGAGTCTTCAGAGCAGCAGGTGATCGGAGCTCGCTCTCTTTGTG 1219  
QY 227 TyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuVal 246  
Db 1220 CAGATCAGTCCATCGCGCGC-----GTGTGCTCATC 1255  
QY 247 AspPheAlaHisValAlaGlu-----GlyAspGly-----ValIleAspHis 260  
Db 1256 GACTTCGGCAAGACCAGCGCCCTCCCGATGGCCAGATCCTGGACCAACCAGCGGCCCTGG 1315  
QY 261 -----AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal 273  
Db 1316 GAGAGGGCAACCGCAGGACGCGTATTTGCTGGGCTGGACAACTCTCATGGCATCTGT 1375  
QY 274 SerAspIleValProGlu-----ThrProHisThrGlnProLeuGlyProSer 289  
Db 1376 GCCAGCTGCTGCTGAGAGATGAGGTGGACTCT-----GTCCCGCGCGCGCTCA 1426

## RESULT 6

US-09-248-796A-5436  
; Sequence 5436, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 5436  
; LENGTH: 963  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-5436

Alignment Scores:

Pred. No.: 2,59e-06 Length: 963  
Score: 139.00 Matches: 67  
Percent Similarity: 37.83% Conservative: 48  
Best Local Similarity: 22.04% Mismatches: 105  
Query Match: 9.14% Indels: 84  
DB: 4 Gaps: 12

US-10-042-894A-8 (1-289) x US-09-248-796A-5436 (1-963)

QY 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAla-----SerAlaSer 19  
Db 223 TCCCAATGACCGCGTTTAAACAACAAGTGGGAGGTATACACCCATATTTCTCAATATCA 282  
QY 20 Lys-----LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGln 36  
Db 283 AAGCAAGAAATATGCAAGAGCTCAACAATACCGAAGAAATTTCTCAAGAAATC--- 339  
QY 37 AlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAla 56  
Db 340 -----TCCAAGAATCATCTTTATTTAT----- 363  
QY 57 ValProAlaArgIleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuPro 76  
Db 364 -----TATATGCCAAGGTATAAAGGTAGTAAT----- 390  
QY 77 ThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAspLeuAlaGly 96  
Db 391 -----GGATCCCAATATTTCTAGAGGATCTTACATCTCAA 426  
QY 97 PheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSer 116  
Db 427 ATGAGAATCTCTGTATATTGGATTTAAAGTGGGACAAGACAATATGGGTGTAATGCA 486  
QY 117 ProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeu 136  
Db 487 ACTATCACCACAAACAATCACAATCAGCTAAGCTAAAGCCAGACGACAAACAAGAAATG 546  
QY 137 GlyPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGlu 156  
Db 547 GGTGTTAGTAATATGCGGACTTCAATCTTTAATATCAAAACAATATTTCTATCAGGAT 606  
QY 157 ArgProGluValLysAlaMetAspThrAlaGlyValArgValLeuArgValTrpVal 176  
Db 607 -----AAATATCTTGAAGAAGACATA 627  
QY 177 SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLys 196  
Db 628 ACTGTTGGTAAACAGTTTGGCAAA-----ATTGTAGCAAAATTTTATATAATGGCCAT 681  
QY 197 Gly-----GlyValLeuSerGlnLeuArgGluLeuLys 207  
Db 682 GACATTTATTCATTATAATCGTATCCCGCATTTAATCGATCACTTAAGAATATATAC 741  
QY 208 AlaTrpPheGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyr 227  
Db 742 ACAATATTACTGGATTACCGGATAGAGTGTATGGTAGTCTGATATTTATTAATGTAT 801  
QY 228 AspAlaAlaValAlaAlaGlyGlyAspGly-----GlyGlyValThrValLys 244  
Db 802 GAA-----GGAGGTGAAGATAAATCTGGAAATCAAGTAAAGTGAAG 843  
QY 245 LeuValAspPheAlaHis---ValAlaGluGlyAspGlyValIleAspHisAsnPheLeu 263  
Db 844 ATTATAGATTTTGTCTAATAGCAGTATAGCGGCGCAGACATATAGATAAT----- 894  
QY 264 GlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThrProHisThr 283  
Db 895 -----GTGACTGTTGCACCACCAACCCCTGATTCC 924  
QY 284 GlnProLeuGly 287  
Db 925 CCCAAACTGGGG 936

## RESULT 7

US-09-252-991A-6525/c  
; Sequence 6525 Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 6525  
; LENGTH: 1899  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6525

Alignment Scores:  
Pred. No.: 0.00663 Length: 1899  
Score: 113.50 Matches: 93  
Percent Similarity: 33.51% Conservative: 35  
Best Local Similarity: 24.35% Mismatches: 115  
Query Match: 7.46% Indels: 139  
DB: 4 Gaps: 20

US-10-042-894A-8 (1-289) x US-09-252-991A-6525 (1-1899)

QY 12 AlaGlyHisArgAlaSerAlaSerLeuGlyProLeuIleAspGlySerGlyLeuPhe 31  
DB 1515 GCGCGCACCGCGTACTTGGTGGCGTGTAGGCGCT-----CATCCGGCGTGGCG 1465  
QY 32 TyrLysProLeuGlnAlaGlyAsp-----ArgGlyGluHisGluVal 45  
DB 1464 GAACAGCCCGAGACCGCTGGAGACATTGACACGTCGGCGCTTCCGGAGGCGCTTCAGGTG 1405  
QY 46 AlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAspThrPhe 65  
DB 1404 CGAAGGAACGCGTGTGGTCGCTTCCACCGCCAGAGAGTTGATGTCATGATCCACATC 1345  
QY 66 PheProArg-----PheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81  
DB 1344 GTATTCGAGTAGTCGCTCTCCACGGTACCGCGTGGCGAGCCCGCGGTGTGTGAA 1285  
QY 82 GlyLupProHisPro-----HisLeuVal 89  
DB 1284 AATCAGTTTACCCGGCCATGTTGCGTGGCGCTTGTGGCCCGGCGCTGCACCTGCTC 1225  
QY 90 -----LeuAspAsp-----LeuLeu 94  
DB 1224 GCGGTGCGGCACGCTGCACGACCGCGGTGGAGACGCGTACCCGCTGCTGGCGAGCAATTG 1165  
QY 95 AlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProPro 114  
DB 1164 GCGGGTTTCTCGAGAGTCGCGCATGACGTGCGCCAGGCGCAGGTGACAAACCTGCT 1105  
QY 115 Ser--SerProGluProTyrIleAlaLysCysLeuAlaLysAspArg----- 129  
DB 1104 GCCAGCTCCAGCGGCGGCGGACCGATGCCGAGCTGCGCGGTGATGGCGCGAC 1045  
QY 130 -----GlyThrThrSer-----ValLeuLeuGlyPheArgVals 141  
DB 1044 TTTGTTTTTGAACGACTTATCGGGGTACCCCTTTCGGCCTGTGGCGAGCTGCTGTTGCTC 985  
QY 141 exGly-----ValArgValValGly----- 147  
DB 984 AGGCGCGGAGCGGTGCTGTTGGTGGTAGCGGAGCGGTGCAACTGCCGGGTGAGG 925  
QY 148 -----ProGluGlyAlaValTrpArgThrGluArgProG 159

DB 924 GCACGGAAGCGCCAGGTGAATCCCGCCACAGGGTGCAGTTGCGGCGCTGACCGGATGC 865  
QY 159 luValLysAlaMetAspThr-----AlaGlyValArgArgValLeuArgA 174  
DB 864 AGGTACCAGCTTGTGCAACCGCCCGCTCCAGACGCTGCGGTCCAGTTCGCGCTGGAGG 805  
QY 174 tGTYrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAla-----AlaAlav 192  
DB 804 TATTCGTTGTAGCGTTCCTGCACCGGGCGCTTGACCTCGAGACTGAGCAGTTGCGGCGC 745  
QY 192 alTyrGly----- 194  
DB 744 TTCATCAGCTTCAGCGCATCGAGGAGTAGGCGATCTGCGACTCGATCATGTAGACCATC 685  
QY 195 -----GlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluG 212  
DB 684 GAGTTGTGCGCGGAGGCGGTGTTGCGTCCCAT-----CAGG 649  
QY 212 luGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAla----- 229  
DB 648 AAGAACAGTTGGGAAGCGCGGTGGTGTAGTGCCTTGTAGGCTTCCGGGCGCTTGGTGC 589  
QY 230 -----AlaAlaValAla-----AlaGlyGlyAsp----- 237  
DB 588 CAGCTGTCCAGCAGGTGCGCGCGCTGCGACCGAAGACCACTCCGCGGGGATAGGATCG 529  
QY 238 --GlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAla----- 252  
DB 528 TTGGCGGTGAAGCGCGGTACCGAAGATGATCGCGTCCGCTCCGCGCGGTGCGCG 469  
QY 253 -----GluGlyAspGlyValIle-----AspHisAsnPheLeuGlyG 265  
DB 468 TCAGCATTCATTGCGCGTGCAGCGCGGATGCTTCGCTGATCACCGTGAGTTGGCG 409  
QY 265 ly 265  
DB 408 GC 407

## RESULT 8

US-09-949-016-2348  
; Sequence 2348, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2348  
; LENGTH: 4496  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-2348

Alignment Scores:  
Pred. No.: 0.0325 Length: 4496  
Score: 112.50 Matches: 59  
Percent Similarity: 34.98% Conservative: 40  
Best Local Similarity: 20.85% Mismatches: 87  
Query Match: 7.40% Indels: 97  
DB: 4 Gaps: 15

US-10-042-894A-8 (1-289) x US-09-949-016-2348 (1-4496)

QY 65 PheProArgPheHisGlyThrArgLeuProThrGluAlaGlnProGlyGluPro 84  
Db 807 TTGTTACCTCCCTACCATGGG-----GATGTGTGAAGGACGGGGAGCGC 851  
QY 85 HisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAsp 104  
Db 852 TACAACCG-----ATGGACGACCTGCTGGCGGACTTCGACTCGCCCTGTGTGATGGAC 905  
QY 105 IleValleGlyAlaIleThrTrp----- 112  
Db 906 TGAAGATGGGAATCAGGACCTACCTGGAGGAGGCTCACGAAGGCCCGGAAGAAGCCCC 965  
QY 113 -----ProProSerPro----- 117  
Db 966 AGCTCGGAGGACATGTACCAAGAGATGATCAGGTGGACCCCGAGGCCCGCCACCGGAG 1025  
QY 118 GluProTyrIleAlaLysCysLeuAlaLysAspArgGly----- 130  
Db 1026 GAGGAAAAAGCACAGCGGGCTGTGACCAAGCCAGGTACATGCGTGGCGGGAGACCATC 1085  
QY 131 ThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValGlyProGluGly 150  
Db 1086 AGCTCCAGCGCCACCTCGGGGTTTCAGGATCGAGGAATCAAG-----AAAGAAGACGGC 1139  
QY 151 AlaValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArg 170  
Db 1140 ACCGTG-----AACCGGGACTTCAGNAGACCCAAAGAGGAGCGAGTCCCGGAG 1190  
QY 171 ValLeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 190  
Db 1191 GCCTTCAGAGAGTTCACTAAA----- 1211  
QY 191 AlaValTyrGlyGlyLysGlyGlyValLeu-----SerGlnLeuArgGluLeu 206  
Db 1212 -----GGAACCAATACATCTCTGATCGCCTATCGGACCGGCTGAAGGCCATT 1259  
QY 207 LysAlaTrpPheGluGlnThrLeuPheHisPheTyr-----SerAlaSerIle 223  
Db 1260 CGAACCACTCTAGAAGTTCTCCCTCTTCAAGTGCACGAGGTATGTCAGCTCCCTC 1319  
QY 224 LeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrVal 243  
Db 1320 CTCTTCATCCACACAGAGGAAACAGGCC-----AAAGTG 1355  
QY 244 LysLeuValAspPheAlaHis-----ValAlaGluGly----- 254  
Db 1356 TGGATGATGACTTTGGGAAAAACACCGCCCTGCTGAGGGCCAGACCTGTCAGCATGAC 1415  
QY 255 -----AspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIle 270  
Db 1416 GTCCCTGGCAGAGGGGAACCGGAGGATGGTACCTCTCGGGGCTCAATAACCTCGTC 1475  
QY 271 LysPheValSerAspIleValProGluThrPro-----HisThrGlnPro 285  
Db 1476 GACATCTGACCGAGATGTCACGAGATGCCCTGCTGCTGAGTGGCCACGCGCCCTCCT 1535  
QY 286 LeuGlyPro 288  
Db 1536 ---GGCCCC 1541

## RESULT 9

US-09-949-016-323  
; Sequence 323, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 323  
; LENGTH: 4505  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-323

Alignment Scores: Pred. No.: 0.0326 Length: 4505  
Score: 112.50 Matches: 59  
Percent Similarity: 34.98% Conservative: 40  
Best Local Similarity: 20.85% Mismatches: 87  
Query Match: 7.40% Indels: 97  
DB: 4 Gaps: 15

US-10-042-894A-8 (1-289) x US-09-949-016-323 (1-4505)

QY 65 PheProArgPheHisGlyThrArgLeuProThrGluAlaGlnProGlyGluPro 84  
Db 810 TTGTTACCTCCCTACCATGGG-----GATGTGTGAAGGACGGGGAGCGC 854  
QY 85 HisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAsp 104  
Db 855 TACAACCG-----ATGGACGACCTGCTGGCGGACTTCGACTCGCCCTGTGTGATGGAC 908  
QY 105 IleValleGlyAlaIleThrTrp----- 112  
Db 909 TGAAGATGGGAATCAGGACCTACCTGGAGGAGGAGCTCACGAAGGCCCGGAAGAAGCCCC 968  
QY 113 -----ProProSerPro----- 117  
Db 969 AGCTCGGAGGACATGTACCAAGAGATGATCGAGGTGGACCCCGAGGCCCGCCACCGAG 1028  
QY 118 GluProTyrIleAlaLysCysLeuAlaLysAspArgGly----- 130  
Db 1029 GAGGAAAAAGCACAGCGGGCTGTGACCAAGCCAGCGGTACATGCGTGGCGGGAGACCATC 1088  
QY 131 ThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValGlyProGluGly 150  
Db 1089 AGCTCCAGCGCCACCTCGGGGTTTCAGGATCGAGGAATCAAG-----AAAGAAGACGGC 1142  
QY 151 AlaValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArg 170  
Db 1143 ACCGTG-----AACCGGGACTTCAAGNAGACCAACCAAGAGGAGGAGGTCCACCGAG 1193  
QY 171 ValLeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 190  
Db 1194 GCCTTCAGAGAGTTCACTAAA----- 1214  
QY 191 AlaValTyrGlyGlyLysGlyGlyValLeu-----SerGlnLeuArgGluLeu 206  
Db 1215 -----GGAACCAATACATCTGATCGCCTATCGGACCGGCTGAAGGCCATT 1262  
QY 207 LysAlaTrpPheGluGlnThrLeuPheHisPheTyr-----SerAlaSerIle 223  
Db 1263 CGAACCACTCTAGAAGTTCTCCCTCTTCAAGTGCACGAGGTGTCATGGCAGCTCCCTC 1322  
QY 224 LeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrVal 243  
Db 1323 CTCTTCATCCACGACAGGAAAGAACAGGCC-----AAAGTG 1358  
QY 244 LysLeuValAspPheAlaHis-----ValAlaGluGly----- 254  
Db 1359 TGGATGATGACTTTGGGAAAAACCAAGCCCTCGCTGAGGGCCAGACCCCTGTCAGCATGAC 1418  
QY 255 -----AspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIle 270

Db 1419 GTCCCTGGCAGGAGCGGGAACCGGAGGATGGCTACCTCTCGGGGCTCAATAACCTCGTC 1478  
Qy 271 LyePheValSerAspIleValProGluThrPro-----HisThrGlnPro 285  
Db 1479 GACATCTGACCGAGATGTCACAGGATGCCCACTGCTGCTGAGTCCCAACGCGCTCCCT 1538  
Qy 286 LeuGlyPro 288  
Db 1539 ---GGCCCC 1544

RESULT 10  
US-09-902-540-1096/c  
; Sequence 1096, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: MYXOCOCCUS xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1096  
; LENGTH: 14555  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(14555)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1096

Alignment Scores:  
Pred. No.: 0.658 Length: 14555  
Score: 108.00 Matches: 81  
Percent Similarity: 34.29% Conservative: 15  
Best Local Similarity: 28.93% Mismatches: 104  
Query Match: 7.10% Indels: 81  
DB: 4 Gaps: 12

US-10-042-894A-8 (1-289) x US-09-902-540-1096 (1-14555)  
Qy 5 HisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeu 24  
Db 2867 CATCGACCTGAAGCGGCTGTTCTCCGCCACGCGGCGACGGAGCCGTCATGCGCGTTG 2808  
Qy 25 Ile-AspGlySerGlyLeu----- 30  
Db 2807 ATCCGATGTCGAGGGGCTGTCGCGGGCTTCGCCGTGAGGCCCAAGAGTGGTCCAGA 2748  
Qy 31 ----PheTyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrG 49  
Db 2747 AGGTACCATGACCTGCTGGAGCTGGACGCGGAGGCTGGAGACGGAGCCCTCACCA 2688  
Qy 49 uAlaPhe-----SerAlaHisAlaAlaValProAlaArgAlaArgAs 63  
Db 2687 AGCTTACCTGCGGCTGGGCGGCCACCTGCACACCTGAAGGCGACGCGCTCGCTGG 2628  
Qy 63 pThrPhePheProArgPheHis-----GlyThrArgLeuLeuProThrGluAl 79  
Db 2627 GCATGACGACCTGAGCGACATGCGCCACAAGCTGGAGGACGCGCTCGCCCGCTGAAGG 2568  
Qy 79 aGlnProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGly----- 96  
Db 2567 CGAACCTCAGAAGATG--CCGCGGCGGTGGTGCATCCTGCTTCACGGCTGGACCT 2510  
Qy 97 -----PheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpPro 114  
Db 2509 CTTCTGCTGCGCGGAGCCCAACGCGACGCGGCTGGGGAGGCGCTTCGCGACCCGCGC 2450

Qy 114 oSerSerProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerVa 134  
Db 2449 GTCCGCCCTGGCGAGCTGGTGGCGAGCCCGCGCGCCCGAGAGGC----- 2401  
Qy 134 lleuLeuGlyPheArgValSerGlyValArgValValGlyPro-----G 149  
Db 2400 -----CGGCCCATGGTCCGCCACCGGGGGCTGGCGCGCTGCCCGCGC 2354  
Qy 149 uGlyAlaValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValAr 169  
Db 2353 CGGTGCGCTC-----CGCGAGCCCGTCTGTCGCGAGCGCGCTGCCCGAGTCCG 2303  
Qy 169 g---ArgValLeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCys---- 186  
Db 2302 CGTACGCGGCTGGCGCGTGTCTGTC-----GTGCGAGTGCAGCGCGCTGATCGCGGAG 2249  
Qy 187 -----AlaLeuAlaAlaValTyrGlyLysGlyValLeuSerGlnLe 203  
Db 2248 TGGAGCGCTGCGGAGGTGCGCTGCGGTGAGGAGCGCGCGCGAACTGGA---- 2194  
Qy 203 uArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSerAlaSeril 223  
Db 2194 ----- 2194  
Qy 223 eLeuLeuGlyTyrAspAlaAla-----AlaValAlaAlaGlyGlyAspGlyGlyGly 240  
Db 2193 -----CGGGTGCGCTGCTGCGGAAGCAGGGGCTGCTGCGGAGCGCGGAGGC 2140

RESULT 11  
US-09-248-796A-4664  
; Sequence 4664, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 4664  
; LENGTH: 954  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-4664

Alignment Scores:  
Pred. No.: 0.0114 Length: 954  
Score: 107.50 Matches: 46  
Percent Similarity: 37.08% Conservative: 20  
Best Local Similarity: 25.84% Mismatches: 53  
Query Match: 7.07% Indels: 59  
DB: 4 Gaps: 9

US-10-042-894A-8 (1-289) x US-09-248-796A-4664 (1-954)  
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 28 TTAATTCACCTACCTCTTCTTAACATCAAGCTGTGCCAT----- 69  
Qy 21 LeuGlyProLeuIleAspGly-----SerGlyLeuPheTyrLysProLeuGln 36  
Db 70 -----GATGGCTGCCTTCAACCGATTTCATTATTATTAAATTGACTGTT 114  
Qy 37 AlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAla 56  
Db 115 -----CAGCAAGAAATTTGATTTCTACACTCAAGTTTCAACTGCAC----- 153



```
QY 57 ValProAlaArgIleArgAspThr-----PhePheProArg 68
Db 154 ---GATCAAGAGTGTCAAGATGCTCCATTAGGTGCCCAATGTCCTGATGCTTACA 210
QY 69 PheHisGlyThr-----Arg 73
Db 211 TTCTGGGACATTGACACAAGGAGATGTATCAAAACTCAAAACAGCAGGTCTGGATCAA 270
QY 74 LeuLeuProThrGluAlaGlnProGlyGluProHisPheHisLeuValLeuAspLeu 93
Db 271 CTGTCATCAGCAGAGACAAAACAGATAAA-----CAATACATTGTACTACTGAATTCG 324
QY 94 LeuAlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAla----- 109
Db 325 TACCATTGGCTTTACCATCTCTAGTATATGGAATCAATCAAGTTGGCGCCAAATTGACTGAC 384
QY 110 -----IleThrTrpProSerProGluProTyrlleAlaLysCysLeuAlaLys 127
Db 385 GATGAAGTTACTGCTCCAGAGAAA-----ATTGCCAGTTTACAAAAAGTC 429
QY 128 AspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145
Db 430 AGTGATTCAACTACTAGTGGAGTTTAAATTTTCAAGATTTTGTGTATGAAAGTC 483

RESULT 12
US-09-252-991A-12872/c
; Sequence 12872, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12872
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12872

Alignment Scores:
Pred. No.: 0.012 Length: 984
Score: 107.50 Matches: 91
Percent Similarity: 35.86% Conservative: 32
Best Local Similarity: 26.53% Mismatches: 117
Query Match: 7.07% Indels: 105
DB: 4 Gaps: 15

US-10-042-894A-8 (1-289) x US-09-252-991A-12872 (1-984)
QY 8 GluHisGlnValAlaGlyHisArgAlaSer-----AlaSerLysLeuGlyPro 23
Db 968 GAAACGCAATGGCGGACGACGCTCTACGTGACCGCCATGGGAATACTGGCGGCGCG 909
QY 24 LeuIleAspGlySerGlyLeuPheTyrlleProLeuGlnAlaGlyAspArgGlyGlu--- 42
Db 908 TCCGCGGACGCGGACGACAGTTT-----ACGTGACGCTGGATTTCCGCGTCTGT 855
QY 43 -----HisGluValAlaPheTyrlleGluAlaPhe----- 51
Db 854 CCACGCTGAAGCAGATTGAGGAGGACGCGCGCTCTCCGCGCTTCCGAGGCTGA 795
QY 52 -----SerAlaHisAlaAlaValPro-AlaArgIleAr 62
Db 794 TCCGCGCGTGGCGACGCGCTGCTCTGCGACGCGGAGCAATTCACGCTGTCTCCG 735
QY 62 AspThr-----PhePheProArgPheHisGlyThrArgLeuLeuProThrGlu 78
```

```
Db 734 GCGTACGCGGCGGCGAGTACCACTTCGCGCGAGCGTGTGGG-----ATGACTGGCG 684
QY 78 ualaGlnProGlyGluProHisPheHisLeuValLeu----- 90
Db 683 AGAACAGCCCGGTGCGCGGCGCGGACCTCTGCTGCTCCAGCGCGCGATACAGGC 624
QY 91 -----AspAspLeuLeuAlaGlyPheGlnAl 99
Db 623 GCGGCTTGGAGCGCGAGATGCTGATGCGGCGCGAGCTTTCTTGGCCGTATAGCAGC 564
QY 99 aProCysValAlaAsp----- 105
Db 563 CCTTCTTGAAGCTGACCCCGCGCGCTGCAAGTTGATCATCTGTGGGATGAACAGCT 504
QY 105 e-LysIleGlyAlaIleThrTrpProSerProGluProTyrlleAlaLysCysVal 125
Db 503 CGCGGTAGCGCGCAACACCTGGCG-----ATGCGTGC-GGCGACTGCGCGAGCAGC 451
QY 125 euAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgV 145
Db 450 CAATCGTCGAGCGCGCTCGCGCTGCTGCTCGC-GGAGCGTGGCAGCACCGCTCGGC 392
QY 145 alValGlyProGlyAlaValTrpArgThrGluArgPro---GluValLysAla-Met 163
Db 391 GCGTTCGCGCGGAACCCATAGCTCGACGCGGCGTCCGCCAGCGCGAGCAGGTC 332
QY 164 AspThrAlaGlyValArgValLeuArgArgTyVal----- 176
Db 331 GCGCTGCGCGGCGATCTCGCGCTCGCGCGGGGTCTATCCCGAGGCGCGCAGTGC 272
QY 177 -----SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaAla 191
Db 271 CTGCTGGCGTCCGCGCAAGCTATGCGAGCCAGGC-----TGGCTCTCTCGTCGCAAG 218
QY 192 ValTyrlleGlyLysGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
Db 217 GCTGGCTTGGAGATACGCGGTATTTCTTCCAGTCCGCAACTGGGCGCTC----- 167
QY 212 GluGlnThrLeuPheHisPheTyrlleAlaSerIleLeuLeuGlyTyrlleAlaAla 231
Db 166 GAGCAGCTCGCGCGCATCGCCAGCAGCGCGCTCGCGCTCGCGGAGGATGCGGAAGCT 107
QY 232 -----ValAlaAlaGlyAspGlyGlyValThrValLys 244
Db 106 CGATAGCAGCGCGCGCTTGTATGTTGAGCGCGCGCGAGCTGGAGTCTCGTCTTGAG 47
QY 245 LeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGly 264
Db 46 -----GTAGCGAGGTTGCAAGTTCAGCTGGCGCTGGAGGAACCTTGGC 5
QY 265 Gly 265
Db 4 GGC 2

RESULT 13
US-09-016-434-1132/c
; Sequence 1132, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
```



```
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 6000 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cdna
US-08-348-006B-6

Alignment Scores:
Pred. No.: 0.251 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 1 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-348-006B-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
DB 3077 AGTCTTGGCGAGACTTGTCCGCGAGAACGTCGCGTAGCGGA----- 3033

QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
DB 3032 ---CGGGGGGCTGAAGGGCCAGGGCCCGCGCGTGTGGCTCGCACTTGGAGGTGTCAT 2976

QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
DB 2975 AGCCCGTGTCCGGCTTCAGGCCCTGCAGCTGAGCGTCT----- 2936

QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
DB 2935 -----CCGCGCCGGCTCAGCGCGCTCGCGGAGTCTAGTCTCTCGGG 2892

QY 82 -----GlyGluProHisProHisLeuValLeuAspLeuLeuAlaG1 96
DB 2891 CAGGGCCAGGCGACCGCGCTCCGCGACGCGCCACCGTGTATTGACGAT----- 2843

QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
DB 2842 -----GGCCCCGTGTGGCTCGCGCGCGCACCGGTGGCAGCGCGGA 2802

QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
DB 2801 GAAGGAC-----GGTCCCGCGCGAGGCGTTCGCCGCGCGCTCCA 2763

QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
DB 2762 GAATCTCGGGTGGCCACGCGGGCGTGTCTCCGCGGATGCTCAGGACCTCGGCTGCTCCT 2703

QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluValLy 161
DB 2702 CGCCCGAGCGCCCGCGGCTCGCGGCGCGCAAGCCGGAACACATAGCTGGCCCGCTTGTGCA 2643

QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
DB 2642 CGCTGTATGCGGTGACCGGTCTTCGAGAGCGGGAACTCCAGGTGGCCAGGGCGGTGCG 2583

QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
DB 2582 AGTCCTCAGCGCCAACTGACGCGGGTAGCCAGCACCTGTCTCCGCGGTGCCAGCGG 2523

QY 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
DB 2522 GGGGCTCCACGCTGCGCAGCAGGCTGCGCTCGGGGCTCTGCTGCACCGA----- 2474

QY 206 uLysAlaTrpPheGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
DB 2473 -CAGGGTGGGGCCCGCACACTGCTCCCTTGTGTCAACACAC-----CT 2430

QY 225 uGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
DB 2429 TGGGTTGTCTCGAGCGCCATCGCCCTTCATGTGTGTAGCGGGCTACCGTGTAGGAGTACG 2370
```

```
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyVal11 258
DB 2369 CGGTCTCAGGCTGCAAGTTTGTGATGACCATCTCTCTGGGATCGGCCAGCATGACGTCCT 2310

QY 258 eAspHisAsnPheLeuGlyGlyLeuCysSerLeuLeuLysPheValSerAspIleValPr 278
DB 2309 TGATGCGCGCGCGCGCGCGCTCCATCGCGCCCTCCATGCGCAGCTAGTGGACCTGGTAGC 2250

QY 278 oGluThrProHisThrGlnProLeuGly 287
DB 2249 CGCGGATCTGGCGCTGCTGCTGCCGCCGGG 2222

RESULT 15
US-08-800-825A-6/c
/ Sequence 6, Application US/08800825A
/ Patent No. 5866397
/ GENERAL INFORMATION:
/ APPLICANT: RODAN, GIDEON A.
/ APPLICANT: SCHMIDT, AZRIEL
/ APPLICANT: RUTLEDGE, SU JANE
/ TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
/ NUMBER OF SEQUENCES: 7
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
/ STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
/ CITY: RAHWAY
/ STATE: NEW JERSEY
/ COUNTRY: USA
/ ZIP: 07065-0900
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/800,825A
/ FILING DATE: 14-FEB-1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: HAND, J. MARK
/ REGISTRATION NUMBER: 36,545
/ REFERENCE/DOCKET NUMBER: 18992DA
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 732-594-3905
/ TELEFAX: 732-594-4720
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 6000 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cdna
US-08-800-825A-6

Alignment Scores:
Pred. No.: 0.251 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 2 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-800-825A-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
DB 3077 AGTCTTGGCGAGACTTGTCCGCGAGAACGTCGCGTAGCGGA----- 3033

QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
DB 3032 ---CGGGGGGCTGAAGGGCCAGGGCCCGCGCGTGTGGCTCGCACTTGGAGGTGTCAT 2976
```

```
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGGCCGTGTCGGGCTTCAGGCCCTTCAGCGTAGCGTTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CCGGCCCGGCTCAGCGGGCTGCCGGCAGCTCAGTCTCTCGGG 2892
QY 82 -----GlyGluProHisProHisLeuValLeuAspLeuAlaGl 96
Db 2891 CAGGGCCAGGCACCGGCTCCCGCAGCCACCGTGTATTGACGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
Db 2842 -----GGCCCCCTTCGCTCGCGGGCAGGGTGGCAGCCAGCGGA 2802
QY 116 rProGluProTyrlleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValleule 136
Db 2801 GAAGGAC-----GGTCCCGCCGAGGGTTGCCGGCGGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValargValVal----- 146
Db 2762 GAATCTGCGGGTGGCCACCGGGGCTGCTCTCGGGATGCTCAGGACCTCGGCTGCCTCT 2703
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161
Db 2702 CGCCCCAGCGCGCGGCTCCGGGCGCAAGCCGGAACACATAGTGGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCCTCATGCCGTGTAGCGGTCTCTCGAGGGCGGGAACTCCAGGTGGCCAGGGGGCTCG 2583
QY 170 gValLeuArgArgTyTyValSerSerValAlaAsp-----GluGlyMetAspCy 186
Db 2582 AGTCCTCAGCGGCAAACTCGAGCGGTAGCCAGCACCTCGTCTCTCCGGGTGCCAGCG 2523
QY 186 sAlaLeuAlaAlaAlaValTyrglyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGCTCCAGCGGTGCCAGAGGCTCCCTCGGGGGTCTGTGACCCGA----- 2474
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyTrSerAlaSerIleLeule 225
Db 2473 -CAGGGTTGGCGGCCAGCACCTGCTCCTTGCTCACAACCAC-----CT 2430
QY 225 uGlyTyTrAspAlaAlaAlaValAlaGlyAspGlyGly----- 239
Db 2429 TGGGTTTGTGCGAGCGCATCGCCCTTCATGTTGTAGGCGGCTACCGTGTGGAGTACG 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValil 258
Db 2369 CGGTCTCAGGCTGCAAGTTTGTGATGACCATCTCTGGGCATCGGCCAGCATGACGTCT 2310
QY 258 eAspHisAsnPheLeuGlyLeuCysSerLeuIleLysPheValSerAspIleValPr 278
Db 2309 TGATGCGCGGCGCGCGGCTCCGGCCCTCCATGCGCAGTAGTGCACCTGGTAGC 2250
QY 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CGCGGATCTGCGCGTCTGCGCGCGG 2222
```

Search completed: June 17, 2005, 06:38:38  
Job time : 230 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2005, 01:52:53 ; Search time 615 Seconds  
(without alignments)  
2781.797 Million cell updates/sec

Title: US-10-042-894A-8  
Perfect score: 1521  
Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHQTPLGPS 289

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Qs/cgn2\_1/USPTO.spool/US10042894/runat\_15062005\_111949\_8281/app\_query.fasta\_1.455  
-DB=N Geneseq 16Dec04 -QFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blastum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=apto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10042894@cgn\_1\_470@runat\_15062005\_111949\_8281 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_16Dec04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID         | Description        |
|------------|--------|-------------|--------|------------|--------------------|
| 1          | 1521   | 100.0       | 1344   | 6 AAD43514 | Aad43514 Maize ino |
| 2          | 1485   | 97.6        | 922    | 6 AAD43513 | Aad43513 Maize ino |
| 3          | 1435   | 94.3        | 923    | 6 AAD43512 | Aad43512 Maize ino |
| 4          | 1406.5 | 92.5        | 1169   | 6 AAD43511 | Aad43511 Maize ino |
| 5          | 1307   | 85.9        | 3416   | 6 AAD43522 | Aad43522 Maize ino |

|    |       |      |      |    |          |                    |
|----|-------|------|------|----|----------|--------------------|
| 6  | 1233  | 81.1 | 899  | 6  | AAD43518 | Aad43518 Maize ino |
| 7  | 808.5 | 53.2 | 643  | 6  | AAD43519 | Aad43519 Maize ino |
| 8  | 742.5 | 48.8 | 1020 | 6  | AAD43517 | Aad43517 P. argent |
| 9  | 722.5 | 47.5 | 1195 | 6  | AAD43516 | Aad43516 Eucalyptu |
| 10 | 714.5 | 47.0 | 1130 | 3  | AAC48750 | Aac48750 Arabidops |
| 11 | 700.5 | 46.1 | 1130 | 3  | AAC40313 | Aac40313 Arabidops |
| 12 | 674.5 | 44.3 | 1104 | 3  | AAC33685 | Aac33685 Arabidops |
| 13 | 674.5 | 44.3 | 1168 | 3  | AAC39023 | Aac39023 Arabidops |
| 14 | 674.5 | 44.3 | 1243 | 3  | AAC38692 | Aac38692 Arabidops |
| 15 | 644.5 | 42.4 | 1105 | 6  | AAD43515 | Aad43515 Soybean i |
| 16 | 484   | 30.5 | 519  | 6  | AAD43520 | Aad43520 Maize ino |
| 17 | 329   | 21.6 | 353  | 6  | AAD43521 | Aad43521 Maize ino |
| 18 | 306.5 | 20.2 | 484  | 10 | ADE82058 | Ade82058 Arabidops |
| 19 | 288   | 18.9 | 2930 | 4  | ABL23306 | Ab123306 Drosophil |
| 20 | 273.5 | 18.0 | 464  | 6  | ABL93575 | Ab193575 Arabidops |
| 21 | 265   | 17.4 | 876  | 4  | ABL23307 | Ab123307 Drosophil |
| 22 | 201.5 | 13.2 | 1251 | 12 | ADK65918 | Adk65918 Human IPM |
| 23 | 201.5 | 13.2 | 2212 | 8  | AAS53176 | Aad53176 Human kin |
| 24 | 193   | 12.7 | 1341 | 11 | ADI31483 | Adi31483 Human cDN |
| 25 | 192   | 12.6 | 2391 | 10 | ADC06742 | Adc06742 Human ino |
| 26 | 192   | 12.6 | 2890 | 6  | AAD38858 | Aad38858 Human kin |
| 27 | 191   | 12.6 | 2608 | 10 | ADA53453 | Ada53453 Human cod |
| 28 | 189   | 12.4 | 1340 | 4  | AAI60607 | Aai60607 Human pol |
| 29 | 189   | 12.4 | 1677 | 13 | ADQ87184 | Adq87184 Human tum |
| 30 | 189   | 12.4 | 1737 | 5  | AAH64793 | Aah64793 Human sec |
| 31 | 189   | 12.4 | 1737 | 10 | ADC06746 | Adc06746 Human ino |
| 32 | 189   | 12.4 | 1739 | 4  | AAH16813 | Aah16813 Human cDN |
| 33 | 189   | 12.4 | 1748 | 4  | AAI58821 | Aai58821 Human pol |
| 34 | 189   | 12.4 | 1748 | 5  | ADQ99042 | Adq99042 DNA encod |
| 35 | 189   | 12.4 | 1748 | 9  | ADB48802 | Adb48802 Novel hum |
| 36 | 189   | 12.4 | 1757 | 5  | AAH64800 | Aah64800 Human sec |
| 37 | 189   | 12.4 | 1791 | 11 | ACN89795 | Acn89795 Breat ca  |
| 38 | 189   | 12.4 | 1990 | 5  | AAH64898 | Aah64898 Human sec |
| 39 | 183   | 12.0 | 1918 | 3  | AACT7195 | Aac77195 Human ORF |
| 40 | 182.5 | 12.0 | 1222 | 10 | ADC06748 | Adc06748 Murine in |
| 41 | 182.5 | 12.0 | 2306 | 10 | ADF44508 | Adf44508 Mouse kin |
| 42 | 182   | 12.0 | 1863 | 8  | ABX70876 | Abx70876 Novel hum |
| 43 | 178.5 | 11.7 | 1893 | 8  | ABX70875 | Abx70875 Novel hum |
| 44 | 168   | 11.0 | 1723 | 10 | ADC06744 | Adc06744 Human ino |
| 45 | 168   | 11.0 | 4461 | 6  | ABK83730 | Abk83730 Human cDN |

ALIGNMENTS

RESULT 1  
AAD43514  
ID AAD43514 standard; DNA; 1344 BP.  
XX  
AC AAD43514;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 52..921  
FT /tag= a  
FT /product= "Maize IPPK protein #4"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 XX WPI; 2002-636540/68.  
 DR P-PSDB; AAE26196.  
 XX  
 XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 XX  
 XX Claim 1; Page 64-65; 86pp; English.  
 XX  
 XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are used to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK DNA  
 XX  
 SQ Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,06e-130 Length: 1344  
 Score: 1521.00 Matches: 289  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43514 (1-1344)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
 DB 52 ATGTCGAGCTCCACCGCGGAGCAACCAAGTCGCGGCACCGCCCTCCGCGAGCAAG 111  
 QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
 DB 112 CTGGGCGCGCTCATCGACGGCTCCGGCTCTTCTACAGCGGCTCCAGGCGGACCGT 171  
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
 DB 172 GGGGAGACGAGGTCCGCTTCTATGAGGCGTTCCTCGCCACGCGCCCTCCGCGCGCC 231  
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
 DB 232 ATCCGAGACACCTTCTTCCCGCGTTCCAGCGCACCGACTCTCCCAACGAGGGCGAG 291  
 QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
 DB 292 CCGGGGAGCGCATCTCCACCTCGCTCTCGACGACTCTCGCGGGTTTCAGGCGGCC 351  
 QY 101 CysValAlaAspIleIleIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
 DB 352 TCGCTCGCAGACATCAAGATCGCGCCATCAGCTGGCCACCGAGTTCGCGGAGCCCTAC 411  
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
 DB 412 ATCGCCAAAGTCTCCGCAAGNCCCGGGGACACAGAGGTTCTGCTCGGATTCGCGTC 471  
 QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
 DB 472 TCCGGGTCCGAGTCGTCGCCCGGAGGCGCGCTGTGGCGAGCGCGGAGGCGGAGGTG 531  
 QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180  
 DB 532 AAGGCCATGGACACCGCGCGTCCCGCGTGTCTCCGCGCTACGTGTCATCCGTTGCC 591  
 QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeu 200

DB 592 GACGAGGGGATGGACTGTGCGCTCGCGCGGCTGTACGGAGGAAAGGTGGAGTCTTG 651  
 QY 201 SerGlnLeuArgGluLeuIlyeAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220  
 DB 652 TCACAGCTCGCGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTGTTCACACTTACTCG 711  
 QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyGly 240  
 DB 712 GGTTCGATTCTTCTGGGCTATGCTGCTGCAGTCCGACAGCGGAGATGGGGTGG 771  
 QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260  
 DB 772 GTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAGGGTGATGGGTGATTGACCCAC 831  
 QY 261 IenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280  
 DB 832 AACTTCTCTGGCGGCTCTGCTCGTGTATCAAGTTCGTTTCTGACATTGTTCCGGAGACT 891  
 QY 281 ProHisThrGlnProLeuGlyProSer 289  
 DB 892 CCTCATACGACGCCCTTTGGGTCTTCT 918

RESULT 2  
 AAD43513  
 ID AAD43513 standard; DNA; 922 BP.  
 XX  
 AC AAD43513;  
 XX  
 DT 14-NOV-2002 (first entry)  
 XX  
 DE Maize inositol polyphosphate kinase (IPPK) DNA #3.  
 XX  
 KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic; gene; ds.  
 XX  
 OS Zea mays.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 53..922  
 FT /\*tag= a  
 FT /product= "Maize IPPK protein #3"  
 XX  
 PN WO200259324-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 09-JAN-2002; 2002WO-US003120.  
 XX  
 PR 12-JAN-2001; 2001US-0261465P.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 XX WPI; 2002-636540/68.  
 DR P-PSDB; AAE26195.  
 XX  
 PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 XX  
 PS Claim 1; Page 62-63; 86pp; English.  
 XX  
 CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are used to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK DNA





Db 173 GGGGAGCAGAGTGGCTTTCTATGAGGGTCTTCGCCCCACGCGCCGCTCCGGCCGC 232  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 233 ATCCGAGACACCTTCTTCCCCCGTTCCACGCGCAGGACTCTCCCCACCGAGGCGAG 292  
Qy 81 ProGlyGluProHisProHisLeuValLeuLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
Db 293 CCGGGGAGCCGATCCGACCTCGCTTCGACGACCTCTCGCGGATTTGAGGCGCC 352  
Qy 101 CysValAlaAspIleValIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 353 TGGGTCCGAGACATCAAGATCGGGCCATCATGTCGCCACCGAGTTCGCCGAGCCCTAC 412  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 413 ATCGCCAAAGTGGCTCGCCATGGACCGCGGACCCACGAGCGTTCTGCTCGGATTCGCGTC 472  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 473 TCGCGCTCCGAGTCGCTCCCGAGGCGCGGTGTGCGGACGAGCGCCCGAGGTG 532  
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAl 180  
Db 533 AAGGCTATGACACCGTCGCGTCCGCCGCGTCTCCGCGGTACGTGTATCCCGTTGC 592  
Qy 180 aAepGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyVal 200  
Db 593 CGACGAGGGATGGACTGCGCGCTCGCGCGCGGTGTACGGAGGAAAAGGTGGAGTCTT 652  
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
Db 653 GTCACTGCTGCGGAGCTCAAGCGGTGTTCGAGGAGCAGCTCTGTTCACCTTCTACTC 712  
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyGlyGlyGly 240  
Db 713 GGGCTCGATTCCTGGGCTATGATCTGCTGCGAGTCGCGAGCGGCGAGGTGGGGTGG 772  
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
Db 773 GGTAAACAGTGAAGCTGGTGACTTTGCCCATGTGGCCGAGGTGATGGGGTGATTGAACA 832  
Qy 260 sAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
Db 833 CAACCTTCCTGGCGGCTCTGCTAGCTGATCAAGTTCTGTTCTGACATTGTTCCAGAGAC 892  
Qy 280 rProHisThrGlnProLeuGlyProSer 289  
Db 893 TCTCTAGACGCGAGCCTTTGGGTCTTCT 920

RESULT 4

ID AAD43511 standard; DNA; 1169 BP.  
XX AAD43511;  
AC AAD43511;  
XX 14-NOV-2002 (first entry)  
DT Maize inositol polyphosphate kinase (IPPK) DNA #1.  
DE Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; db.  
XX Zea mays.  
XX Key Location/Qualifiers  
FT CDS 84..806  
FT /\*tag= a  
FT /product= "Maize IPPK protein #1"  
XX WO200259324-A2.  
PN  
XX

PD 01-AUG-2002.  
XX 09-JAN-2002; 2002WO-US003120.  
XX 12-JAN-2001; 2001US-0261465P.  
XX (PION-) PIONEER HI-BRED INT INC.  
PA Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
PI P-PSDB; AAE26193.  
DR WPI; 2002-636540/68.  
XX P-PSDB; AAE26193.  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX Claim 1; Page 58-59; 86pp; English.  
XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
SQ Sequence 1169 BP; 238 A; 362 C; 347 G; 222 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.29e-120 Length: 1169  
Score: 1406.50 Matches: 277  
Percent Similarity: 96.21% Conservative: 2  
Best Local Similarity: 95.52% Mismatches: 10  
Query Match: 92.47% Indels: 2  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43511 (1-1169)

Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
Db 84 ATGCCCGACCTCCACCCCGCGAGCACCAGTCGCGGTACCGCCCTCCCGCAGCAAG 143  
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
Db 144 CTGGGCCCGCTCATCGACGGCTCCGCGCTCTTCTTACAAGCCGCTCCAGGCCGCGACCGT 203  
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
Db 204 GGGGAGCAGAGTGGCTCTTATGAGGCGTTCTCGCCACGCGCGCTCCCGGCCCGC 263  
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
Db 264 ATCCGAGACACCTTCTTCCCCCGTTTCCACGCGACCGGACTCTCTCCACCGAGGCGAG 323  
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 324 CCGGGGAGCGCATCCGCACCTCTCTCTCGACGACCTCTCTCGCGGGTTTGAGGCGCC 383  
Qy 101 CysValAlaAspIleValIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120  
Db 384 TCGGTCCGAGACATCAAGATCGCGCCATCAGCTGCGCCACCGAGTTCGCGCGAGCCCTAC 443  
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 444 ATCCGCAAGTACTCTGCCAAGGACCGCGGACCAAGAGGTTCTGCTCGGATTCGCGTCC 503  
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
Db 504 TTGC--GTCCGAGTCTGTCGGCCCCGAGGCGCGGTGTGGCGGACGAGCGCCGAGGTG 561



QY 161 LysAlaMetAspThrAlaGlyValArgAtqValLeuArgArgTyrValSerSer-ValAl 180  
 DB 562 AAGGCTATGGACACCGTCGGCGGTCCGCGGTGCTCCGGCGCTACGTCATCCGCTGC 621  
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValle 200  
 DB 622 CGACGAGGGATGAGCTGCGCGCTCGCGCGCGGTGTACGGAGGAAAAGTGGAGTCIT 681  
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
 DB 682 GTCACAGCTCGCGAGCTCAAGCATGTTGGAGGAGCAGACTCTGTTCACCTTCTACTC 741  
 QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGl 240  
 DB 742 GCGTTCGATTCCTTCTGGGCTATGATGCTGTCAGTCGCACGAGCGAGGTGGGGTGG 801  
 QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
 DB 802 GGTAAACAGTGAAGCTGGTGACTTTGCCCATGTGGCCGAGGGGTGATGGGGTGATTGACCA 861  
 QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280  
 DB 862 CAACCTTCCTGGGCGAGCTCTGCTAGCTGATCAAGTTTCGTTCTGCACATTGTCAGAGAC 921  
 QY 280 rProHisThrGlnProLeuGlyProSer 289  
 DB 922 TCCTTAGACGACGACCTTTGGGTCTCTTCT 949

## RESULT 5

ID AAD43522 standard; DNA; 3416 BP.

XX AAD43522;

DT 14-NOV-2002 (first entry)

DE Maize inositol polyphosphate kinase (IPPK) DNA #9.

KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic; gene; ds.  
 OS Zea mays.

FH Key Location/Qualifiers

FT CDS 72..407

FT FT /\*tag= a

FT FT /product= "Maize IPPK protein #6"

XX W0200259324-A2.

XX 01-AUG-2002.

XX 09-JAN-2002; 2002MO-US003120.

XX 12-JAN-2001; 2001US-0261465P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX WPI; 2002-636540/68.

XX P-PSDB; AAE26201.

XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
 PT increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.

XX Claim 1; Page 75-77; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful

CC in modulating the phytic acid biosynthesis by decreasing phytate and/or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK DNA

XX Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.9e-110 Length: 3416  
 Score: 1307.00 Matches: 260  
 Percent Similarity: 90.34% Conservative: 2  
 Best Local Similarity: 89.66% Mismatches: 9  
 Query Match: 85.93% Indels: 20  
 DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AAD43522 (1-3416)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
 DB 72 ATGCCGAGCTTCACCCGCGGAGCACCAGTCGCGGTCCCGGCTCCCGCAGCAG 131  
 QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
 DB 132 CTGGGCGCACTCATCGACGACTCTGGCTCTTCTACAGCGCTCCAGGCGCGGACCGT 191  
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
 DB 192 GGGGAGCAGAGTCGCTTCTATGAGGGGTTCTCGGCCACCGCCGCTCCGCGGCCGC 251  
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
 DB 252 ATCCGAGACACCTTCTTCCCGGTTCCACGGCAGCGACTCTCTCCACCGAGGCGCAG 311  
 QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100  
 DB 312 CCGGGGAGCGCATCCGACCTCTCTCGACGACTCTCTCGGGGTTTGAGGCGCCC 371  
 QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
 DB 372 TGGTCCGAGACATCAAGATCGGTGCCATCACGTG----- 406  
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
 DB 407 -----ACCACGAGCGTCTCTCGGATTCGCGGTC 436  
 QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160  
 DB 437 TCGGGGTCCGAGTCGTCCGCCCGGAGGGCGCGGTGTGGGGAGGAGCGCCCGGAGGTG 496  
 QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180  
 DB 497 AAGGCTATGGACATTCGCGCGTCCCGCGCTCTCGCGGCTACGTCGTATCGCTTGC 556  
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValle 200  
 DB 557 CGACGAGGGGTGGAGTGGCGCTCGCGCGCGGTGTACGGAGGAAAAGTGGAGTCIT 616  
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220  
 DB 617 GTCACAGCTCGCGAGCTCAAGCGGTGGTTTCGAGGGGCGAGACTCTGTTCACCTTCTACTC 676  
 QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyGl 240  
 DB 677 GGGTTCGATTCCTTCTGGGCTATGATGCTGTCAGTCGCAGTCGCAGCAGGCGGAGGTGG 736  
 QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260  
 DB 737 GGTAAACAGTGAAGCTGCTGACTTTGCCCATGTGGCGAGGGGTGATGGGGTGATTGACCA 796  
 QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280

```

Db      797 CAACCTTCCTGGCGGCTCTGCTAGCTGATCAAGTTTGTTCGACATTGTTCCAGAGAC 856
QY      280 rProHisThrGlnProLeuGlyProSer 289
Db      857 TCCTCAGACGCGCCTTTGGTCCCTTCT 884

RESULT 6
AAD43518
ID      AAD43518 standard; DNA; 899 BP.
AC      AAD43518;
XX      14-NOV-2002 (first entry)
DT      Maize inositol polyphosphate kinase (IPPK) DNA #5.
DE      Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW      nutritional value; animal feed; transgenic; gene; ds.
XX      Zea mays.
FH      Key
FT      Location/Qualifiers
CDS      89..424
          /*tag= a
          /product= "Maize IPPK protein #4"

WO200259324-A2.
01-AUG-2002.
09-JAN-2002; 2002WO-US003120.
12-JAN-2001; 2001US-0261465P.
(PION-) PIONEER HI-BRED INT INC.
Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
WPI; 2002-636540/68.
P-PSDB; AAE26200.

New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.

Claim 1; Page 73-74; 86pp; English.
CC      The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA

XX      Sequence 899 BP; 140 A; 312 C; 276 G; 171 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2,44e-104      Length:      899
Score:          1233.00      Matches:      248
Percent Similarity: 88.42%      Conservative: 4
Best Local Similarity: 87.02%      Mismatches: 14
Query Match:    81.07%      Indels:      20
DB:            Gaps:      1

US-10-042-894A-8 (1-289) x AAD43518 (1-899)

QY      1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20

```

```

Db      89 ATGCCCGACCTCCACCCCGCGGAGCACCAAGTCGCGGTTCACCCGCGCTCCGCCAGCAAG 148
QY      21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db      149 CTGGGCCCACTCATCGAGCGGTCTGGCCTCTTTACAAGCCGCTCCAGCCCGCGCACCGT 208
QY      41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db      209 GGGGAGACGAGGTGCGCTTCTATAGGGGTCTCCGCCACGCGCTCCGCCGCCGCC 268
QY      61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db      269 ATCCGAGACACCTTCTCCCGCGTTCCAGCGCAGCGACTCTCTCCACCGAGCGCAG 328
QY      81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db      329 CCCGGGAGCGCATCCGTACCTCTCTCGACGACCTCTCTCGGGGTTTGAGGGGCC 388
QY      101 CysValAlaAspIleLysIleGlyAlaIleThrTTPProProSerSerProGluProTyr 120
Db      389 TCGTTCGAGACATCAAGATCGGTGCCATCAGTGACCATG-AGCAT----- 435
QY      121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db      436 -----CTGCTCGATTCCACGTC 453
QY      141 SerGlyValArgValValGlyProGluGlyAlaValTTPArgThrGluArgProGluVal 160
Db      454 TCCGCGCTCCGAGTCGTCCGCCCGAGGGCGCGGTGTGGCGAGCGAGCGCTCAGGTG 513
QY      161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
Db      514 AAGGCTATGGACATTTGTCGGGTCCCGCGCTGCTCCGGCGTGCATGTCTCCGTTGC 573
QY      180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValVal 200
Db      574 CGCGAGGGGATGGACTCGCGCTCGCGCGCGGTGTACGGAGGAGAAAGGTGGAGTCTT 633
QY      200 uSerGlnLeuArgGluLeuLysAlaTTPheGluGlnThrLeuPheHisPheTyrSe 220
Db      634 GTCACAGCTGCGCGAGCTCAAGCGCTGGTTTCAGGGGCGAGACTCTGTTCACCTTACTC 693
QY      220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db      694 GCGTTCGATTCTTCGGGCTATGATGCTGCTGCAGTCGACGAGCGAGGTGGGGTGG 753
QY      240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db      754 GGTAACAGTGAAGCTGGTGGACCTTGGCCCATGTGGCGAGGGGTGATGGGGTGAATGACCA 813
QY      260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db      814 CAACCTTCCTGGCGGGCTCTGCTAGCTGATCAAGTTTGTTCGACATTGTTCCAGAGAC 873
QY      280 rProHisThrGln 284
Db      874 TCCTTAGACGCGAG 886

RESULT 7
AAD43519
ID      AAD43519 standard; DNA; 643 BP.
XX      AAD43519;
AC      AAD43519;
XX      14-NOV-2002 (first entry)
DT      Maize inositol polyphosphate kinase (IPPK) DNA #6.
DE      Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW      nutritional value; animal feed; transgenic; ds.
XX      Zea mays.

```

PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PIONEER) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
XX WPI; 2002-636540/68.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 74; 86pp; English.  
XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
SQ Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;  
  
Alignment Scores:  
Pred. No.: 2,41e-65 Length: 643  
Score: 808.50 Matches: 162  
Percent Similarity: 93.14% Conservative: 1  
Best Local Similarity: 92.57% Mismatches: 10  
Query Match: 53.16% Indels: 3  
DB: 6 Gaps: 0  
  
US-10-042-894A-8 (1-289) x AAD43519 (1-643)  
QY 1 MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20  
DB 117 ATGCGGACCTCCACCGCGCGAGCACCAAGTCGCGGTCCCGCGCTCCGCGAGCAAG 176  
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40  
DB 177 CTGGGCGCGCTATCGACGCTCGGCGCTCTTACAGCGCTCCAGGCGCGGACCGT 236  
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60  
DB 237 GGGGAGCAGGAGTGGCTTCTATAGGCGTTCCTCGCCACCGCGCGCGCGCGC 296  
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80  
DB 297 ATCCGAGACACTTCTTCCCGCGGTTCACGCGACGCGACTCTCCCGACCGGCGCAG 356  
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100  
DB 357 CCCGGGAGCGGATCCGACCTCGCTCTCGACACTCTCTCGCGGTTTGAGCGCCC 416  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120  
DB 417 TCGTTCGACATCAAGATCGGCGCCATCAGGTGGCCACCGAGTTCGCGGAGCCCTAC 476  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
DB 477 ATCGNCAAGTACCTNCGCAAGACCGCGGACCGAGCGTCTCTCGATTCCCGGTC 536  
QY 141 SerGlyValArgValValGlyProGluGlyValValTrpArgThrGluArg-ProGluVa 160

DB 537 TTGC-GTCCGAGTCGTGCGCCCGAGGCGCGGTGTGGCGGACGCGCCCGGGGT 594  
QY 160 llysalametaspthr-AlaGlyValArgValLeuArg 173  
DB 595 GAANGCTATGGACACCCGTCGCGNGNCCGCGGNGTGTCTCGG 635  
  
RESULT 8  
AAD43517  
ID AAD43517 standard; DNA; 1020 BP.  
XX  
AC AAD43517;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.  
XX  
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;  
KW nutritional value; animal feed; gene; ds.  
XX  
OS Parthenium argentatum.  
XX  
FH Key Location/Qualifiers  
FT CDS 21..908  
FT /tag= a  
FT /product= "P. argentatum IPPK protein #1"  
XX  
PN WO200259324-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 09-JAN-2002; 2002WO-US003120.  
XX  
PR 12-JAN-2001; 2001US-0261465P.  
XX  
PA (PIONEER) PIONEER HI-BRED INT INC.  
XX  
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
XX WPI; 2002-636540/68.  
DR P-PSDB; AAR26199.  
XX  
XX New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 71-72; 86pp; English.  
XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is P. argentatum IPPK DNA  
XX  
SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 5,26e-59 Length: 1020  
Score: 742.50 Matches: 149  
Percent Similarity: 67.83% Conservative: 45  
Best Local Similarity: 52.10% Mismatches: 79  
Query Match: 48.82% Indels: 13  
DB: 6 Gaps: 5  
  
US-10-042-894A-8 (1-289) x AAD43517 (1-1020)  
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23  
DB 24 CTCAGCGCCCGACATCATCAGTTTCTGGACATGAAGCTGGGCTCGGGAAGCTTGGCCCA 83

```
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 84 CTCATTGATGATTCAGCGCGGTTTACAAACACATGCGAGGTGATTAACCGTGGTCAGNA 143
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
DB 144 GAAGTAGCCTTTTATGAAATCTTTCTTCTAACATAATATTCAGAACACATACGC--- 200
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
DB 201 AAATTCCTTCTATATATTATGACCAAAATCATG-----AAGCATCCACTGGCTCT 254
QY 84 ProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysValAla 103
DB 255 GACCATCTCCATCATGGTGTTCAGATCTTATCATCATGCTCATGTCNACCATCTGTATG 314
QY 104 AspIleLeuIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
DB 315 GACATCAAAATCGGTCCAGAACATGGCGCCAGAGGCTTCCGAGCGGTACATTGCAAAA 374
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
DB 375 TGCCTTAAAAAAGATAGGGAAGACCAAGTATTCCATTGGGATTCAGGATCTCCGGCTG 434
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
DB 435 CAAGTCTATATCGATGATGGTCAGCGGTTTATAGCCTCATAGAAATACATGGGTAAA 494
QY 163 MetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAlaAspGlu 182
DB 495 ACCGGCCAGCTGATGTAGACTACTTCTTAGGAAATTTGTTCTTCTAACCGCTGTGCA 554
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
DB 555 GAGATGGAATGCGCACAGCGCTAGGCCCGGATTTCTTTAGCATCTTTTGTATGGT 614
QY 195 GlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThr 214
DB 615 GGGCCTAATGGGATATTAGCTCACTGATGGAAATGAGACATGGTTTGAAATCAACA 674
QY 215 LeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAla 234
DB 675 ATTTACCACTTCTATGCTTGTCTTTTGTTCATCTTTGAAAGAGGTTGGTTAAAA 734
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly 254
DB 735 GGTGCT---CGGTCAACGACAGAGTCAAACTATTGATTTTGCTCATCTTACAGATGT 791
QY 255 AspGlyValIleAspHisAenPheLeuGlyGlyLeuCysSerLeuIleLysPheValSer 274
DB 792 AATGGTGTATTGATCACAATTTCTTGGTGGGCTCTGTCTTTGATAAAGTTCAATCT 851
QY 275 AspIleValProGluThr 280
DB 852 GACATACTTTCGGAGACA 869
RESULT 9
AAD43516
ID AAD43516 standard; DNA; 1195 BP.
XX AC
AC AAD43516;
XX AC
DT 14-NOV-2002 (first entry)
XX DE
DE Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.
XX KW
KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX KW
KW nutritional value; animal feed; gene; ds.
XX OS
OS Eucalyptus grandis.
XX FH
FH Key Location/Qualifiers
```

```
FT CDS 116..1048
FT /*tag= a
FT /product= "E. grandis IPPK protein"
XX
PN WO200259324-A2.
PD 01-AUG-2002.
PF 09-JAN-2002; 2002WO-US003120.
PR 12-JAN-2001; 2001US-026146SP.
PA (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafaleki JA, Cahoon RE;
XX WPI; 2002-636540/68.
DR P-PSDB; AAE26198.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 68-70; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is Eucalyptus grandis IPPK DNA
XX
SQ Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 4,498-57 Length: 1195
Score: 722.50 Matches: 149
Percent Similarity: 66.31% Conservative: 38
Best Local Similarity: 52.84% Mismatches: 84
Query Match: 47.50% Indels: 11
DB: 6 Gaps: 6
```

```
US-10-042-894A-8 (1-289) x AAD43516 (1-1195)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
DB 119 CTCAAGGTCCCGATCATCAAGTCGCGGTACCCGGGAGACGGGGAAAGCTGGGGCA 178
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 179 CTGGTGGATGATTCCGGCGCGCTTCTATAAGCCTCTCCAGAGCGCATCCGGAGACAG 238
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
DB 239 GAAATGGCCTTTTACGAGTCATTCTTCCAAATACCGAGATCCAGGTCCACATTCGC--- 295
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
DB 296 AAATTCCTTCTCGCGTTCACGGAACCTAAGACTATT-----GAGCGTCTGATGGATCG 349
QY 84 ---ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
DB 350 GGTCTCTCAACCTCACCTGGTTCCTGGAGATCTCGTCTCGGTCGCGACGAAACCATCTCTC 409
QY 103 AlaAspIleLeuIleGlyValaIleThrTrpProProSerSerProGluProTyrIleAla 122
DB 410 ATGGACATCAAGACTGGATCCAGAACATGGTATTCGGAGGCTCTGAGGAGTACATCCAA 469
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
```











XX AAC33685; PR 18-JUN-1999; 99US-0139462P.  
AC 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
DT 17-OCT-2000 (first entry) PR 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
DE 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
KW 23-JUN-1999; 99US-0140354P.  
KW 24-JUN-1999; 99US-0140695P.  
KW 24-JUN-1999; 99US-0140823P.  
XX 28-JUN-1999; 99US-0140991P.  
OS 29-JUN-1999; 99US-0140991P.  
XX 30-JUN-1999; 99US-0141287P.  
PN 01-JUL-1999; 99US-0141842P.  
XX 01-JUL-1999; 99US-0142154P.  
XX 02-JUL-1999; 99US-0142055P.  
PD 06-JUL-1999; 99US-0142390P.  
XX 08-JUL-1999; 99US-0142803P.  
XX 09-JUL-1999; 99US-0142920P.  
XX 12-JUL-1999; 99US-0142977P.  
XX 13-JUL-1999; 99US-0143542P.  
XX 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS EP1033405-A2.  
PN 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.

```

PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

```

## Alignment Scores:

```

Pred. No.: 11e-52 Length: 1104
Score: 674.50 Matches: 141
Percent Similarity: 65.36% Conservative: 42
Best Local Similarity: 50.36% Mismatches: 86
Query Match: 44.35% Indels: 11
DB: 3 Gaps: 7

```

US-10-042-894A-8 (1-289) x AAC33685 (1-1104)

```

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 132 CTCAAAGTCCCTGAACATCAGGTTCAGGACACATTGCTAAAGACGGGAAGCCTGGTCT 191
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 192 CTCGTAGATGACAAAGGTCGGTTCCTCAAGCCACTTCAGGGCGATTCCTCGTGGTGAATC 251
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAep 63

```

```

Db 252 GAGGTAAAGTCTACGAATCTTTCTCTCAAAACAGAGGTTCCAGAACACATCCAT--- 308
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 309 AGATATTTCCCGGTGTATCAGGCACCTCAAGCAGTT-----GAAGGTTCTGATGGAGCA 362
QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 363 -----GCCATGATGGTGTGGAAATCTTCTGCAATACTCAAAACCATCAGTAATG 416
QY 104 AspIleLysIleGlyValAlaIleThrProProSerSerProGluProTyrIleAlaLys 123
Db 417 GATGTTAAGATGGGTTCCGAGAACATGGTATCTGATGCATCTGAGAGATTCATCAATCCAAA 476
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 477 TGTTTGAAGAAAGACACGGGTACCAACCGGTGTATCGGGTTTCAGGATCTCTGGTTTC 536
QY 144 ArgValValGly---ProGluGlyAlaValTTPArgThrGluArgProGluValLysAla 162
Db 537 GAAGTGTATGATCAAAAGAAATCTGGAAGCCCGAGAGGAAGCTTCTTCGCGGG 596
QY 163 MetAspThrAlaGlyValArgValLeuArgArgTyrValSer-----SerValAla 180
Db 597 CTCGATGTAGATGGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCTAACTCACTTTTCG 656
QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGly 198
Db 657 GACACTGGCTCGAAACCTGACTCTGCTTTTGGCTCGAGTGTTCAGCGGGTTCACACGGG 716
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGluGlnThrLeuPheHisPhe 218
Db 717 ATCTTAACGCAAGTCTGCTGGAACCTCAAGACCTGGTTCGAGAACCAACCGCTTACCATTTC 776
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGly 238
Db 777 AACTCTTGTTCGATTTTAATGGTCTATGAGATGAATCCATCTTGAAGGGAATGATGAT 836
QY 239 GlyGly---ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 837 GATGCTAGACCAACAAGCTCAAGCTGGTGGATTTTGTCTCATGTTCTTGTGATGATGATGTC 896
QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuLeuLysPheValSerAspIleVal 277
Db 897 ATTGACCATAACTTCTTGGGTGGTCTTTGCTCTTTCATAAACTTCATCGTGAGATTCIT 956

RESULT 13
AAC39023
ID AAC39023 standard; DNA; 1168 BP.
XX AAC39023;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23077.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23077.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.

```

PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.



PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139859P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0134768P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148119P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 24-SEP-1999; 99US-0155486P.  
PR 28-SEP-1999; 99US-0155659P.  
PR 29-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157753P.  
PR 07-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

```
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 28-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,27e-52 Length: 1243
Score: 674.50 Matches: 141
Percent Similarity: 65.36% Conservative: 42
Best Local Similarity: 50.36% Mismatches: 86
Query Match: 44.35% Indels: 11
DB: 3 Gaps: 7

US-10-042-894A-8 (1-289) x AAC38692 (1-1243)
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLeuGlyPro 23
Db 269 CTCAAAGTCCTGACATCAGGTTCAGGACACATTGCTAAAGCGGAAGCTCGTCT 328
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 329 CTCGTAGATGACAAAGGTCGGTTCCTCAAGCCACTTCAGGGCGATTCTCGTGGTGAATC 388
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAsp 63
Db 389 GAGGTAAAGTTCACGAATCTTCTCTCAACACAGAGGTTCCAGAACACATCCAT--- 445
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 446 AGATATTTCCGGTGTATCATCGGCATCTCAGCAGTT-----GAAGTCTGTGAGGCA 499
QY 84 ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 500 -----GCCATGATGGTGTGGAAATCTTCTTGCAGAACTACTCAAAACCATCAGTAATG 553
QY 104 AspIleGlyIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
Db 554 GATGTTAAAGTGGTTCGAAACATCGGTATCTCGATGCATCTGAAGAATACATCCAAAA 613
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 614 TGTTTGAAGAAACACCGGTACACACACCGTCTCATCGGTTTCAGGATCTCGTTTC 673
QY 144 ArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLysAla 162
Db 674 GAAGTGTATGATCACAAAGAATCGAGTTTCTGGAAGCCGAGAGGAAGCTTCTTCGCGG 733
QY 163 MetAspThrAlaGlyValArgArgValLeuAtqArgTyrValSer-----SerValAla 180
Db 734 CTCGATGTAGATGGAGCGAGATTGATCTCGAGGAAGTTTGTATCATCTAACTCACTTCG 793
QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGly 198
Db 794 GACACTGGCTCGAAACCTGACTCTGCTTTGCTCGAGTGTTCACGCGGTCCACGGG 853
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPhe 218
Db 854 ATCTTAACGCGATGTCTGGAACCTCAAGACCTGGTTCGAGAACCAACGCTCTACCATTC 913
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyAspGly 238
Db 914 AACTCTGTGCTGAATTTAATGGTCTATGAGAAATGAATCCATCTTGAAGGGAATGATGAT 973
QY 239 GlyGly---ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 974 GATGCTAGACCACCAAGTCAAGCTGGTGGATTTTGCTCATGTTCTTGATGGTAAATGGTCT 1033
QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 1034 ATTGACCATAACTTCTTTGGGTGGTCTTTTGTCTTTTCATAAACTTCATCTCGTGAGATTCTT 1093
```

```
RESULT 15
AAD43515
ID AAD43515 standard; DNA; 1105 BP.
XX
AC AAD43515;
DT 14-NOV-2002 (first entry)
XX
DE Soybean inositol polyphosphate kinase (IPPK) DNA.
XX
KW Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 12..851
FT FT /*tag= a
FT FT /product= "Soybean IPPK protein"
XX
PN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
XX
DR P-PSDB; AAE26197.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 66-67; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is soybean IPPK DNA
XX
SQ Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 6,43e-50 Length: 1105
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservative: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 6 Gaps: 8
```

```
US-10-042-894A-8 (1-289) x AAD43515 (1-1105)
QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerLysLeuGlyProLeuIleAsp 26
Db 24 CCGGAGCACCGAGTGGCGGGCACAAGCCAGGACGGAATCTCGGGCCCACTCGTCGAC 83
QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42
Db 84 GATTTTGGAAAATTTACAAGCCCTCCAGACCAACAAAGACGACGACACCGCGGCTCC 143
```



```
Qy 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArg 62
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 ACCGAACCTCTCTTTACACTCTCTCGCGCC---GCCGCCACAGACTACTCATCCGC 200
Qy 63 AspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly 82
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 ---TCCTTCTTCCCGCTTTCACGCGACCGCTCTCTGGACGCTCCGACGGCTCCGGT 257
Qy 83 GluProHisProHisLeuAlaPheValLeuAlaGlyPheGlnAlaProCysVal 102
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
258 ---CCCCACCTCACCTGGCTCTGGAGGACTCTCTGCGGCTACTCCAAACCTCCGTC 314
Qy 103 AlaAspIleValIleGlyAlaIleThrTrpProSerSerProGluProTyrIleAla 122
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 ATGACGTAAGATCGCTCCAGAACCTGGACCTGGGAGACTCCGAGGACTACATCTGC 374
Qy 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
375 AAGTGCCTGAAGAAGGACAGAGACTCTCTAGCTTGGCTTGGGTTTCAGAAATCTCGGA 434
Qy 143 ValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAla 162
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
435 GTCAAG-----GACTTATCTCTCTCTGGGAACCTACCAAGGAAATCTCTCCAGTG 485
Qy 163 MetAspThrAlaGlyValArgValLeuArgValTyrValSerSer----- 178
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
486 CTATCCGCCCATGGTGTGCTGCTTGTCTCAACAGTTCGTTCTCTAATAATATCAAC 545
Qy 179 ValAlaAspGluGlyMetAspCysAlaLeuAlaAlaAlaValTyrGlyLysGlyGly 198
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
546 CATGATCATCATCCGATTGGCTTTCGCAACGGAGGTCTAC-----GGCGCC 596
Qy 199 ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPhe 218
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
597 GTTTGGAGCGCTTGCAGAAAGCTCAAGGACTGGTTCGAGGTTTCAGACGGTGTATCACTTC 656
Qy 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGly 238
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 TATCTTGTCTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 710
Qy 239 GlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIle 258
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
711 ACCAACCTCTGTGTCAAACTCGTTGACTTTTGACACGCTGGTGGACGGAAACGGTGTCA 770
Qy 259 AspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
771 GATCAAACTTCTTGGGTGGCCTTTGTCTTCTCATCAAGTTCTCTCAAGGATATCCTA 827
```

Search completed: June 17, 2005, 04:11:12  
Job time : 625 secs

**This Page Blank (uspto)**



VAFYEAFSAHAAPARIDTFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFAQPCV  
ADIKGAIWTPSSPPDYIAKCLAKDRGTTSVLLGFRVSGVRVVEGAVWTERPEV  
KAMDYAGVRRLRYVSSVADEGMDCALAAAVYGGKGVLSQLRELKAWFERQTLPHF  
YSASILLGVDAAVAAGSGGGVTVKLVDFAHVAEGDGVTDHNFLLGGLSLIKFVSDI  
VPETPTQPLGPS"

ORIGIN

Alignment Scores:  
Pred. No.: 9,21e-106 Length: 1344  
Score: 1521.00 Matches: 289  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AX513570 (1-1344)

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys     | 20  |
| DB | 52  | ATGTCGCGACTCCACCGCGGAGCACCAAGTCGCGCGCCACCGCGCTCCCGCCAGCAAG    | 111 |
| QY | 21  | LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg  | 40  |
| DB | 112 | CTGGGCGCGCTCATCGACGCTCGGCGCTTCTACAAGCGCTCCAGCGCGGCGACCGT      | 171 |
| QY | 41  | GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg     | 60  |
| DB | 172 | GGGAGACAGAGGTGCGCTTCTATGAGCGTTCCTCGCGCCACGCGCGCTCCCGCGCGC     | 231 |
| QY | 61  | IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln  | 80  |
| DB | 232 | ATCCGAGACACCTTCTTCCCGCGTTCCACGGCACCGGACTCTCTCCACCGAGGCGCAG    | 291 |
| QY | 81  | ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro  | 100 |
| DB | 292 | CCGGGGAGCGCATCTCACCTCGCTCTCGACGACTCTCTCGCGGGTTTCAGGGCGCC      | 351 |
| QY | 101 | CysValAlaAspIleValIleGlyAlaIleThrTyrProProSerSerProGluProTyr  | 120 |
| DB | 352 | TGCGTCCGACACATCAAGATCGCGCGCATCACTGCGCCACCGAGTTTCGCGGAGCCCTAC  | 411 |
| QY | 121 | IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal  | 140 |
| DB | 412 | ATGCCCAAGTGCTCCGCAAGACCGCGGAGCACGAGGTTCTGCTCGGATTCGCGTTC      | 471 |
| QY | 141 | SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal  | 160 |
| DB | 472 | TCCGGCGTCCGAGTCTCGCGCCCGAGGCGCGTGTGCGGACGCGCGCGAGGTG          | 531 |
| QY | 161 | LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla  | 180 |
| DB | 532 | AAGGCCATGGACACCGCGCGGTTCGCCGCGTCTCCGCGCTACGTGTCTATCCGTTGCC    | 591 |
| QY | 181 | AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyValGlyGlyValLeu     | 200 |
| DB | 592 | GACGAGGGATGGACTGTGCGCTCGCGCGGGGTGTACGGAGAAAGAGTGAGTCTTG       | 651 |
| QY | 201 | SerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSer  | 220 |
| DB | 652 | TCACAGCTGCGGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCG    | 711 |
| QY | 221 | AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyClyAspGlyGly     | 240 |
| DB | 712 | GGGTGCAATCTTCTGGGCTATGATGCTGTCAGTCGACGAGCGCGAGATGGGGTGGG      | 771 |
| QY | 241 | ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis  | 260 |
| DB | 772 | GTGACCGTGAAGCTGTGTGACTTTTGCCCATGTGGCCAGGGGTGATGGGGTGATTGACCAC | 831 |
| QY | 261 | AsnPheLeuGlyLysSerLeuIleLysPheValSerAspIleValProGluThr        | 280 |
| DB | 832 | AACTTCTCGGGCGGCTCTGCTCGTGAATCAAGTTCGTTCTGACATTGTTCCGGAGACT    | 891 |

|    |     |                             |     |
|----|-----|-----------------------------|-----|
| QY | 281 | ProHisThrGlnProLeuGlyProSer | 289 |
| DB | 892 | CCTCATACGACGCTTTGGTCTCTTCT  | 918 |

RESULT 2

|            |  |             |     |        |                 |
|------------|--|-------------|-----|--------|-----------------|
| AX513568   | AX513568   | 923 bp      | DNA | linear | PAT 05-OCT-2002 |
| LOCUS      | Sequence 5 from Patent WO02059324.                                 |             |     |        |                 |
| DEFINITION | AX513568   |             |     |        |                 |
| ACCESSION  | AX513568.1   | GI:23559668 |     |        |                 |
| VERSION    |  |             |     |        |                 |
| KEYWORDS   |  |             |     |        |                 |
| SOURCE     | Zea mays   |             |     |        |                 |
| ORGANISM   | Zea mays   |             |     |        |                 |
|            | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |             |     |        |                 |
|            | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  |             |     |        |                 |
|            | clade; Panicoideae; Andropogoneae; Zea.                            |             |     |        |                 |

|           |   |   |
|-----------|---|---|
| REFERENCE | 1 | Shi, J., Beach, L.R., Wang, H., Rafaleki, J.A. and Cahoon, R.E. |
| AUTHORS   |   | Novel inositol polyphosphate kinase genes and uses thereof      |
| TITLE     |   | Patent: WI 02059324-A 5 01-AUG-2002;                            |
| JOURNAL   |   | PIONEER HI-BRED INTERNATIONAL, INC. (US)                        |

|          |  |
|----------|--|
| FEATURES | Location/Qualifiers                                      |
| source   | 1..923   |
|          | /organism="Zea mays"                                     |
|          | /mol_type="unassigned DNA"                               |
|          | /db_xref="taxon:4577"                                    |
|          | 53..922  |
|          | /note="unnamed protein product"                          |
|          | /codon_start=1   |
|          | /protein_id="CAD52953.1"                                 |
|          | /db_xref="GI:23559669"                                   |
|          | /translation="MPDLHPHEHQVAGHRASAKPGPLIDGSLFYKPLQAGDRGEHE |

|     |  |
|-----|--|
| CDS |  |
|     | VAFYEFSAHAAPARIDTFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFAQPCV     |
|     | ADIKGAIWTPSSPPDYIAKCLAKDRGTTSVLLGFRVSGVRVVEGAVWTERPEV    |
|     | KAMDYAGVRRLRYVSSVADEGMDCALAAAVYGGKGVLSQLRELKAWFERQTLPHF  |
|     | YSASILLGVDAAVAAGSGGGVTVKLVDFAHVAEGDGVTDHNFLLGGLSLIKFVSDI |
|     | VPETPTQPLGPS"  |

ORIGIN

Alignment Scores:  
Pred. No.: 3.1e-103 Length: 923  
Score: 1485.00 Matches: 283  
Percent Similarity: 98.27% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 5  
Query Match: 97.63% Indels: 0  
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AX513568 (1-923)

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys    | 20  |
| DB | 53  | ATGCCGAGACTCCACCGCGGAGCACCAAGTCGCGCGCTACCGCGCTCCCGCCAGCAAG   | 112 |
| QY | 21  | LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg | 40  |
| DB | 113 | CCGGGCGCGCTCATCGACGCTCCGCGCTTCTTCTACAAGCGCTCCAGCGCGGCGACCGT  | 172 |
| QY | 41  | GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg | 60  |
| DB | 173 | GGGAGACGAGGTTCGCTTCTATGAGCGTTCCTCCGCCACGCGCGCTCCCGCGCGC      | 232 |
| QY | 61  | IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln | 80  |
| DB | 233 | ATCCGAGACACCTTCTCCCGCGTTCCACGCGACGCGACTCTCTCCACCGAGGCGCAG    | 292 |
| QY | 81  | ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro | 100 |
| DB | 293 | CCCGGGAGCGCATCCGACCTCTCTCGACGACCTCTCTCGCGGATTCGAGGCGCC       | 352 |
| QY | 101 | CysValAlaAspIleLysIleGlyAlaIleThrTyrProProSerSerProGluProTyr | 120 |
| DB | 353 | TGCGTCCGAGACATCAAGATCGGCGCATCACTGCGGCGACCGAGTTCGCGGAGCCCTAC  | 412 |

|                   |   |   |                            |
|-------------------|---|---|----------------------------|
| Qy                | 121   | IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal    | 140                        |
| Db                | 413   | ATCGCCAAGTGCCTCGCCATGAGCCGCGGGACCAACGAGCGCTTCGTCTCGAATTCGGGTC   | 472                        |
| Qy                | 141   | SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluAurProGluVal    | 160                        |
| Db                | 473   | TCGGCGCTCGAGTCGTGCGCGCCGAGGGCGCGTGTGGCGGACGAGCGCCCGAGGTG        | 532                        |
| Qy                | 161   | LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla       | 180                        |
| Db                | 533   | AAGGCCATGACACCGCGCGCGTCCGCGGTGTCTCCGGCGCTACGTGTTCATCGTTGCC      | 592                        |
| Qy                | 181   | AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeu          | 200                        |
| Db                | 593   | GACGAGGGATGACTGTGCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTG        | 652                        |
| Qy                | 201   | SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer    | 220                        |
| Db                | 653   | TCACAGCTGCGCGAGCTCAAGGGCGTGGTTCGAGGAGCAGACTCTGTGTCCACTCTCTACTCG | 712                        |
| Qy                | 221   | AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyGly    | 240                        |
| Db                | 713   | GGCTCGATTCTTCGGCTATGATGCTGCTGCAATGTCGACGAGCGGAGGTGGGGTGGG       | 772                        |
| Qy                | 241   | ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis    | 260                        |
| Db                | 773   | GTACCGTGAAGCTGTGGACTTTGCCATGTGGCCGAGGCGTGTATGGGTGATTGACCAC      | 832                        |
| Qy                | 261   | AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr    | 280                        |
| Db                | 833   | AACTTCTCGCGCGGCTCTGCTCGCTGATCAAGTTCTGTCATTTGTTCAGAGACT          | 892                        |
| Qy                | 281   | ProHisThrGlnProLeuGlyProSer                                     | 289                        |
| Db                | 893   | CCTCAGACGCGCTTTGGGTCTCTCT                                       | 919                        |
| RESULT 3          |   |   |                            |
| LOCUS             | AX513566  | 923 bp  | DNA linear PAT 05-OCT-2002 |
| DEFINITION        | Sequence 3 from Patent WO02059324.                              |   |                            |
| ACCESSION         | AX513566  |   |                            |
| VERSION           | AX513566.1  | GI:23559666   |                            |
| KEYWORDS          |   |   |                            |
| SOURCE            | Zea mays  |   |                            |
| ORGANISM          | Zea mays  |   |                            |
| REFERENCE         | 1   |   |                            |
| AUTHORS           | Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.B. |   |                            |
| TITLE             | Novel inositol polyphosphate kinase genes and uses thereof      |   |                            |
| JOURNAL           | Patent: WO 02059324-A 3 01-AUG-2002;                            |   |                            |
| FEATURES          | PIONEER HI-BRED INTERNATIONAL, INC. (US)                        |   |                            |
| source            | Location/Qualifiers   |   |                            |
|                   | 1..923  |   |                            |
|                   | /organism="Zea mays"  |   |                            |
|                   | /mol_type="unassigned DNA"                                      |   |                            |
|                   | /db_xref="taxon:4577"   |   |                            |
| CDS               | 53..736   |   |                            |
|                   | /note="unnamed protein product"                                 |   |                            |
|                   | /codon_start=1  |   |                            |
|                   | /protein_id="CAD52952.1"  |   |                            |
|                   | /db_xref="GI:23559667"  |   |                            |
|                   | /translation="MPDLHPPEHQVAGHRASAKPGLIDGSLFYKPLQAGDRGEHE         |   |                            |
|                   | VAFYFAHAAPARIRDTFFPRPHGTRLLPTEAQGPHEPHLVLDLDAFEAPCV             |   |                            |
|                   | ADIKIGATWPPSPSPYIAKCLAMDRGTTSLIGFRVSGVRVVPBGAVWRTERREV          |   |                            |
|                   | KAMDTGVRLVRLRYVSSACRGGDLGRAGGVRKRKRWLVTAARAQGVVRAASVPL          |   |                            |
|                   | LLGVDSGL"   |   |                            |
| ORIGIN            |   |   |                            |
| Alignment Scores: |   |   |                            |
| Pred. No.:        | 1.85e-99  | Length:   | 923                        |
| Score:            | 1435.00   | Matches:  | 277                        |

|   |                                    |   |                            |
|---|------------------------------------|---|----------------------------|
| Percent Similarity: 96.21%                  |                                    | Conservative: 2   |                            |
| Best Local Similarity: 95.52%               |                                    | Mismatches: 10  |                            |
| Query Match: 94.35%                         |                                    | Indels: 1   |                            |
| DB: 6                                       |                                    | Gaps: 0   |                            |
| US-10-042-894A-8 (1-289) x AX513566 (1-923) |                                    |   |                            |
| Qy  | 1                                  | MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys     | 20                         |
| Db  | 53                                 | ATGCCCGAGACTCCACCGCGGAGACCAAGTCGCGGTCAACCGCGCTCCGCCAGCAAG     | 112                        |
| Qy  | 21                                 | LeuGlyProLeuIleaspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg  | 40                         |
| Db  | 113                                | CCGGCGCGCGCTCATCGACGCGTCCGGCGCTTCTTACAAGCGCTCCAGGCGCGGACCGT   | 172                        |
| Qy  | 41                                 | GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg     | 60                         |
| Db  | 173                                | GGGGAGCAGAGTTCGCTTCTATGAGGCGTCTCCGCCACCGCGCGCTCCCGCGCGC       | 232                        |
| Qy  | 61                                 | IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln  | 80                         |
| Db  | 233                                | ATCCGAGACACTTCTTCGCGCGGTTCACGGCACCGGACTCTCTCCCGCCACGAGGCGCAG  | 292                        |
| Qy  | 81                                 | ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro     | 100                        |
| Db  | 293                                | CCCGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCGCGGATTTGAGGCGGCC        | 352                        |
| Qy  | 101                                | CysValAlaAspIleLysIleGlyAlaIleThrTyrProSerProGluProTyr        | 120                        |
| Db  | 353                                | TGGCTCGCAGACATCAAGATCGCGCCATCACGTGGCCACCGAGTTCCCGGAGCCCTAC    | 412                        |
| Qy  | 121                                | IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal  | 140                        |
| Db  | 413                                | ATCGCAAGTGCCTCGCCATGGACCGCGGACCAACGAGCGTTCTGCTCGATTCCGGCTC    | 472                        |
| Qy  | 141                                | SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluAurProGluVal  | 160                        |
| Db  | 473                                | TCGGCGTCCGAGTCTGTCGTCGCGGCGCGTGTGGCGGACGAGCGCCCGAGGTG         | 532                        |
| Qy  | 161                                | LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl  | 180                        |
| Db  | 533                                | AAGGCTATGGACACCTCGCGCTCCCGCGGTCTCCGGCGCTACGTTGTCATCCGCTGC     | 592                        |
| Qy  | 180                                | aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLe        | 200                        |
| Db  | 593                                | CGACGAGGGGATGCATCGCGCTCGCGCGCGGTGTACGGAGGAGGAGGAGTCTT         | 652                        |
| Qy  | 200                                | uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe  | 220                        |
| Db  | 653                                | GTCACTGCTGCGGAGCTCAAGGCGTGGTTCGAGGAGCAGCCTCTGTTCACCTTCTACTC   | 712                        |
| Qy  | 220                                | rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly       | 240                        |
| Db  | 713                                | GGCGTCGATTCTTCTGGGCTATGATGCTGCTCAGTCCGACGAGCGCGGAGGTGGGGTGG   | 772                        |
| Qy  | 240                                | yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi  | 260                        |
| Db  | 773                                | GGTAACAGTGAAGCTGGTGACTTTGCCCATGTGGCGAGGTGATGGGTGATTGACCA      | 832                        |
| Qy  | 260                                | sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr | 280                        |
| Db  | 833                                | CAACTTCTCGCGCGGCTCTGCTAGCTGATCAAGTTCGTTTCTGACATTGTTCCAGAGAC   | 892                        |
| Qy  | 280                                | rProHisThrGlnProLeuGlyProSer                                  | 289                        |
| Db  | 893                                | TCCTCAGACGAGCCTTTGGGTCTCTCT                                   | 920                        |
| RESULT 4                                    |                                    |   |                            |
| LOCUS                                       | AX513564                           | 1169 bp   | DNA linear PAT 05-OCT-2002 |
| DEFINITION                                  | Sequence 1 from Patent WO02059324. |   |                            |
| ACCESSION                                   | AX513564                           |   |                            |
| VERSION                                     | AX513564.1                         | GI:23559664   |                            |

```

KEYWORDS
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1
AUTHORS     Shi,J., Beach,L.R., Wang,H., Rafaleki,J.A. and Cahoon,R.E.
TITLE       Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL     PATENT: WO 02059324-A 1 01-AUG-2002;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source      1..1169
            /organism="Zea mays"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4577"
            84..806
            /notes="unnamed protein product"
            /codon_start=1
            /protein_id="CAD52951.1"
            /db_xref="GI:23559665"
            /translation="MPDLHPPEHQVAGHRASAKLGPLIDGSLFYKPLQAGDRGEHE
            VAFYAFSAHAAPVIRIDTFPPRHGTRLLPTEAQPGPHLVLDLLAGFEAPCV
            ADIKGAIWPPSSPEPIYAKYAKDRGTTVLLGPRVLRPSRRRGRGVADGAPGE
            GYGHRRPRPRALRVLRLPTRGWTARSRRCTEEKVESCHSCASSRHGWRSLCSTS
            TRRRFWAMLLQSQAEVGVG"
ORIGIN
Alignment Scores:
Pred. No.:      3,46e-97      Length:      1169
Score:          1406.50      Matches:    277
Percent Similarity: 96.21%      Conservative: 2
Best Local Similarity: 95.52%      Mismatches: 10
Query Match:    92.47%      Indels:      2
DB:             6            Gaps:        0

US-10-042-894A-8 (1-289) x AX513564 (1-1169)

QY      1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB      84 ATGCCGACCTCCACCGCGGAGCACCAGTCGCGCGTCCACGCGCTCCGCGAGCAAG 143
QY      21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB      144 CTGGGCGCGCTCATCGAGCGCTCCGCGCTCTTCTACAGCGCTCCAGCGCGGCGACCGT 203
QY      41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB      204 GGGGAGCAGAGTCGCTTCTATGAGCGTTCTCGCCACGCGCTCCGCGCGCGC 263
QY      61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB      264 ATCCGAGACACCTTCTTCCCGCGTTCCAGCGACGCGACTCTCTCCACCGCGCGCAG 323
QY      81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB      324 CCGGGGAGCGGATCCGACCTCGTCTCGAGACTCTCTCGGGGTTTGGGCGCCC 383
QY      101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB      384 TGGCTCGCAGACATCAAGATCGCGCCATCAGTCGCGCCACCGAGTTCCGCGGAGCCCTAC 443
QY      121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB      444 ATCGCAAGTACCTCGCCCAAGGACCGCGGACCCAGCAGCGTTCTGCTCGGATTCGCGCTC 503
QY      141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
DB      504 TTGC--GTCCGAGTCGTGCGGCCCGGAGGCGCGCTGTGGCGGACGAGCGCGCGAGGTG 561
QY      161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB      562 AAGGCTATGACACCGCTCGGCGTCCGCGCTGCTCCGCGCTACGTGTCATCCGCTGCG 621

```

---

```

QY      180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLe 200
DB      622 CCACGAGGGATGGACTCGCGCTCGCGCGCGGTGTACGAGGAGGAGGAGGAGTCTT 681
QY      200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSe 220
DB      682 GTCAAGCTGCGCGAGCTCAGGATGTTGGAGGAGGAGACTCTGTTCCATCTTACTTC 741
QY      220 rLaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB      742 GCGCTCGATTCTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
QY      240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValLeAspHi 260
DB      802 GGTACAGTGAAGCTGGTGGACTTTGCCCATGTGCGCGGAGGTGATGGGTTGATTGACCA 861
QY      260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
DB      862 CAACCTCTCTGGCGGAGCTCTGCTAGTGTATGATCAATGCTTCTGACATTTGTTCCAGAG 921
QY      280 rProHisThrGlnProLeuGlyProSer 289
DB      922 TCCTTAGACGCGAGCCTTTGGGTCCTTCT 949
RESULT 5
AX513583
LOCUS      AX513583
DEFINITION Sequence 20 from Patent WO02059324.
ACCESSION AX513583
VERSION    AX513583.1 GI:23559684
KEYWORDS   Zea mays
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1
AUTHORS     Shi,J., Beach,L.R., Wang,H., Rafaleki,J.A. and Cahoon,R.E.
TITLE       Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL     PATENT: WO 02059324-A 20 01-AUG-2002;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source      1..3416
            /organism="Zea mays"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4577"
            72..407
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD52959.1"
            /db_xref="GI:23559685"
            /translation="MPDLHPPEHQVAGHRASAKLGPLIDGSLFYKPLQAGDRGEHE
            VAFYAFSAHAAPVIRIDTFPPRHGTRLLPTEAQPGPHLVLDLLAGFEAPCV
            ADIKGAIW"
CDS
            1..3416
            /organism="Zea mays"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4577"
            72..407
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD52959.1"
            /db_xref="GI:23559685"
            /translation="MPDLHPPEHQVAGHRASAKLGPLIDGSLFYKPLQAGDRGEHE
            VAFYAFSAHAAPVIRIDTFPPRHGTRLLPTEAQPGPHLVLDLLAGFEAPCV
            ADIKGAIW"
ORIGIN
Alignment Scores:
Pred. No.:      3,95e-89      Length:      3416
Score:          1307.00      Matches:    260
Percent Similarity: 90.34%      Conservative: 2
Best Local Similarity: 89.66%      Mismatches: 9
Query Match:    85.93%      Indels:      20
DB:             6            Gaps:        1

US-10-042-894A-8 (1-289) x AX513583 (1-3416)

QY      1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB      72 ATGCCGACCTCCACCGCGGAGCACCAGTCGCGCGTCCACGCGCTCCGCGAGCAAG 131
QY      21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40

```

```
Db 132 CTGGGCCCATCTACGACGACTCTGGCCCTCTTCTACAAGCGCTCCAGCGCGCGACCGT 191
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 192 GGGAGACAGAGGTGGCTTCTATAGAGCGTTCTCGGCCACCGCGCTCCGGCCCGC 251
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 252 ATCCGAGACACCTTCTCCCGCGTCCACGCGACGCGACTCTCCACCGAGGCGCAG 311
Qy 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
Db 312 CCGGGGAGCGCATCGCACCTCGTCTCGACGACCTCTCGCGGGTTTGAGGGGCC 371
Qy 101 CysValAlaAspIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 372 TGGGTGCGACACATCAAGATCGGTGCCATCAGTG----- 406
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgVal 140
Db 407 -----ACCACGAGCGTTCTGCTCGGATTCGCGTC 436
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 437 TCGCGGTCTCGAGTCTGCGGCCCGAGGCGCGTGTGGCGGACGAGCGCGGAGGTG 496
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 497 AAGGCTATGGACATTGTGCGCGTCCGCGCGTCTCGCGCGTACGTTGTCGCTGC 556
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200
Db 557 CGACGAGGGAGTACGATGCGCGTCTCGCGCGCGGTGTACGAGGAGAAAGTGGAGTCTT 616
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 617 GTCACAGCTGCGCGAGCTCAAGCGTGTTCGAGGGGCGACACTCTGTTCCACTTCTACTC 676
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
Db 677 GGGCTCGATTCTTCTGGGCTATGATGCTGTCAGTGCAGTCGACGAGCGGAGGTGG 736
Qy 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 737 GGTAAACAGTAGCTGTGAGCTTTGCCCATGTGGCGAGGTGTATGGGGTGAITGACCA 796
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 797 CAACTTCTGGCGGGCTCTGCTAGCTGATCAAGTTGTTGTTCTTGACATTGTTCCAGAGAC 856
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 857 TCCTCAGACGACGCTTTGGTGCCTTCT 884

RESULT 6
AX513578 899 bp DNA linear PAT 05-OCT-2002
LOCUS
DEFINITION Sequence 15 from Patent WO02059324.
ACCESSION AX513578
VERSION AX513578.1 GI:23559679
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
Shi,J., Beach,L.R., Wang,H., Rafaleki,J.A. and Cahoon,R.E.
AUTHORS Novel inositol polyphosphate kinase genes and uses thereof
TITLE Patent: WO 02059324-A 15 01-AUG-2002;
JOURNAL PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..899
```

```

/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
89..424
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52958.1"
/db_xref="GI:23559680"
/translation="MPDLHPPEHVOAGHRASAKLGLIDGSLFYPKPLQAGDRGEHE
VAFYFAFSAAHAAVPAIRIDTFFPRFHGTRLLPTEAQGPBHPHVLVLDLLAGFEAPCV
ADIKIGAIT"

CDS
ORIGIN
Alignment Scores:
Pred. No.: 3-21e-84 Length: 899
Score: 1233.00 Matches: 248
Percent Similarity: 88.42% Conservative: 4
Best Local Similarity: 87.02% Mismatches: 14
Query Match: 81.07% Indels: 20
DB: 6 Gaps: 1
US-10-042-894A-8 (1-289) x AX513578 (1-899)
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 89 ATGCCCGACCTCCACC CGCGGAGCACCAAGTCGCGGTCCACCGCTCCGCGCAGCAAG 148
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 149 CTGGGCCCACTCATCGACGCGCTCTGCGCTCTTCTACAAGCGCTCCAGCGCGGCGCGT 208
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 209 GGGAGACAGAGTCCGCTTCTATAGAGCGTCTCCGCCACCGCGCTCCGCGCGCGC 268
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 269 ATCCGAGACACCTTCTTCCCGCGTTCACGCGACGCGACTCTCTCCACCGAGGCGCAG 328
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db 329 CCGGGGAGCGCGATCGTACCTCGTCTCGACGACTCTCTCGCGGGTTTGAGGCGCCC 388
Qy 101 CysValAlaAspIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 389 TGGGTGCGACACATCAAGATCGGTGCCATCAGTGCATG-AGCGAT----- 435
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgVal 140
Db 436 -----CTGCTCGGATTCCACGTC 453
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 454 TCGCGGTCCGAGTCGTCGCGCCCGAGGCGCGGTGTGGCGGACGAGCGCCCTGAGGTG 513
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 514 AAGCTATGACACATTGTCGCGCTCCCGCGCTCTCGGCGCTGCATGTCATTCGCTTGC 573
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLe 200
Db 574 CGCGGAGGGATGAGTCTGCGCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGAGTCTT 633
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 634 GTCACAGCTGCGGAGCTCAAGCGTGTGTTTCGAGGGGCGAGACTCTGTTCACCTTCTACTC 693
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyLysAspGlyGlyGln 240
Db 694 GGGTTCGATTCTTCTGGGCTATGATGCTGTCAGTGCAGCAGCGAGCGGAGGTGG 753
Qy 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 753 -----
```



|            |   |   |            |                 |
|------------|---|---|------------|-----------------|
| Db         | 754   | GGTAACAGTGAAGCTGGTCGACCTTCCCATGTGCGGAGGTGATGGGTGATTGACCA      | 813        |                 |
| Qy         | 260   | sAnpHeLeuGlyGlyLeuCySerLeuIleIysPheValSerAspIleValProGluTh    | 280        |                 |
| Db         | 814   | CAACTTCCTCGGGCGGCTCTGCTAGCTGATGATCAAGTTTGTTCGTACATGTTCCAGAGAC | 873        |                 |
| Qy         | 280   | rProHisThrGln   | 284        |                 |
| Db         | 874   | TCCTTAGACGCAG   | 886        |                 |
| RESULT 7   | AP005749  | 149142 bp   | DNA linear | PLN 29-JUN-2004 |
| LOCUS      | Oryza sativa (japonica cultivar-group)  | genomic DNA, chromosome 2,                                    |            |                 |
| DEFINITION | BAC clone:OSJNBa0047A17.  |   |            |                 |
| ACCESSION  | AP005749  |   |            |                 |
| VERSION    | AP005749.3  | GI:49388911   |            |                 |
| KEYWORDS   | Oryza sativa (japonica cultivar-group)  |   |            |                 |
| SOURCE     | Oryza sativa (japonica cultivar-group)  |   |            |                 |
| ORGANISM   | Oryza sativa (japonica cultivar-group)  |   |            |                 |
| REFERENCE  | Sasakawa, T., Matsumoto, T. and Katayose, Y.  |   |            |                 |
| AUTHORS    | Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  |   |            |                 |
| TITLE      | Clone:OSJNBa0047A17   |   |            |                 |
| JOURNAL    | Published Only in Database (2002)   |   |            |                 |
| REFERENCE  | 2 (bases 1 to 149142)   |   |            |                 |
| AUTHORS    | Sasakawa, T., Matsumoto, T. and Katayose, Y.  |   |            |                 |
| TITLE      | Direct Submission   |   |            |                 |
| JOURNAL    | Submitted (18-SEP-2002) Takuji Sasakawa, National Institute of  |   |            |                 |
|            | Agrobiological Sciences, Rice Genome Research Program; Kamondai   |   |            |                 |
|            | 2-1-2, Tsukuba, Ibaraki 305-8602, Japan   |   |            |                 |
|            | (E-mail:tsasakawa@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,  |   |            |                 |
|            | Tel:81-298-38-7441, Fax:81-298-38-7468)   |   |            |                 |
| COMMENT    | On Jun 28, 2004 this sequence version replaced gi:34740257.   |   |            |                 |
|            | Genes were predicted from the integrated results of the following:  |   |            |                 |
|            | GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH  |   |            |                 |
|            | (http://www.softberry.com/), GeneMark.hmm   |   |            |                 |
|            | (http://opal.biology.gatech.edu/GeneMark/), GlimmerM  |   |            |                 |
|            | (http://www.tigr.org/tcdb/glimmer/glmr form.html), RiceHMM  |   |            |                 |
|            | (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor  |   |            |                 |
|            | (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4  |   |            |                 |
|            | (http://globin.cse.psu.edu/html/docs/sim4.html), gap2   |   |            |                 |
|            | (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The   |   |            |                 |
|            | genomic sequence was searched against NCBI NonRedundant Protein   |   |            |                 |
|            | database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA   |   |            |                 |
|            | sequence database at RGP or DDBJ. Protein homologies of the coding  |   |            |                 |
|            | regions were searched against NCBI NonRedundant Protein database  |   |            |                 |
|            | with BLASTP. ESTs represent the identified cDNA sequences using   |   |            |                 |
|            | BLASTN with the corresponding DDBJ accession no. and RGP clone ID.  |   |            |                 |
|            | Full-length cDNAs represent the identified cDNA sequences using   |   |            |                 |
|            | BLASTN with the corresponding DDBJ accession no.  |   |            |                 |
|            | A gene with identity or significant homology to a protein is  |   |            |                 |
|            | classified based on the protein name to indicate the homology level   |   |            |                 |
|            | such as same name, 'putative-' and '-like protein'. A gene without  |   |            |                 |
|            | significant homology to any protein but with full-length cDNA or  |   |            |                 |
|            | EST homology (covering almost the entire length of partial  |   |            |                 |
|            | sequence) is classified as an 'unknown' protein. A gene predicted   |   |            |                 |
|            | by two or more gene prediction programs is classified as a  |   |            |                 |
|            | 'hypothetical' protein according to IRGSP standard. A gene  |   |            |                 |
|            | predicted by a single gene prediction program is also classified as   |   |            |                 |
|            | a probable 'hypothetical' protein and is included as a  |   |            |                 |
|            | miscellaneous feature of the sequence.  |   |            |                 |
|            | The orientation of the sequence is from -21M13 to M13rev of the BAC   |   |            |                 |
|            | clone. This sequence of OSJNBa0047A17 clone has an overlap with   |   |            |                 |
|            | OSJNB003322 clone (DDBJ: AP005842) at 5' end and an overlap with  |   |            |                 |
|            | P0415B12 (DDBJ: AF004772) at 3' end. Detailed information on  |   |            |                 |
|            | overlap and assembly quality together with annotation of this entry   |   |            |                 |
|            | is available at <a href="http://rgp.dna.affrc.go.jp/GenomeSeq.html">http://rgp.dna.affrc.go.jp/GenomeSeq.html</a> . |   |            |                 |
| FEATURES   | Location/Qualifiers   |   |            |                 |
| source     | 1. .149142  |   |            |                 |

/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="OSJNBa0047A17"  
join(2637..2780,3298..3360)  
/gene="OSJNBa0047A17.1"  
join(2637..2780,3298..3360)  
/gene="OSJNBa0047A17.1"  
/note="hypothetical ORF  
predicted by GENSCAN  
this category is not included in IRGSP standard"  
complement(5864..9107)  
/gene="OSJNBa0047A17.2"  
complement(join(5864..6755,8843..9107))  
/gene="OSJNBa0047A17.2"  
/note="supported by full-length cDNA(s): AK102815"  
complement(join(6112..6755,8843..9017))  
/gene="OSJNBa0047A17.2"  
/note="Rieske iron-sulfur protein  
contains EST(s): AU165272(C63256),AU068962(C51210)  
contains full-length cDNA(s): AK102815"  
/codon\_start=1  
/product="putative ubiquinol-cytochrome c reductase  
iron-sulfur subunit, mitochondrial precursor"  
/protein\_id="BAD26134.1"  
/db\_xref="GI:49388912"  
/translation="MLRVRRLLTTALWRPAAAGAGPLAGSLPGDDEFSPQR  
PRFVDSPPFAASRGSETLVRNQDVLTELPAIVSAVKNPSSAKIVVDEYNHRYQ  
PGDPSKRAPFVLSGGRTFVLSLLRLVLFVLSMSAKVLALASLEVLSSIEPG  
TTVTVKRGKPVFIRRTTEEDINLANSVDIGSLRDPQDAERVKNPWLVLVGVCTHL  
GCILPNAAGDGGWFCPCGSHYDISGRKRPAPYNLEVPYFSLEENKLLIG"  
9904..10750  
/gene="OSJNBa0047A17.3"  
join(9904..10067,10732..10750)  
/gene="OSJNBa0047A17.3"  
/note="start and end point are not identified"  
join(9904..10067,10732..10750)  
/gene="OSJNBa0047A17.3"  
/note="predicted by GENSCAN etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD26135.1"  
/db\_xref="GI:49388913"  
/translation="MENATVTADLAPWSQQCRPVGEEREGVAAAKKEGRQIQKE  
EKGGRRRSGGIRKTM"  
complement(11599..12180)  
/gene="OSJNBa0047A17.4"  
complement(<11599..>12180)  
/gene="OSJNBa0047A17.4"  
/note="supported by full-length cDNA(s): AK109226"  
complement(11599..12180)  
/gene="OSJNBa0047A17.4"  
/note="contains full-length cDNA(s): AK109226  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
join(15870..16043,16479..16484)  
/gene="OSJNBa0047A17.5"  
join(15870..16043,16479..16484)  
/gene="OSJNBa0047A17.5"  
/note="hypothetical ORF  
predicted by FGENESH  
this category is not included in IRGSP standard"  
complement(18932..25995)  
/gene="OSJNBa0047A17.6"  
complement(join(18932..19648,19740..19836,20081..20194,  
20280..20444,20690..20827,20923..20999,21093..21207,  
21277..21382,21703..21764,21859..21993,22149..22331,  
22880..22936,23369..23417,23504..23652,24724..24771,  
25583..25995))  
/gene="OSJNBa0047A17.6"  
/note="supported by full-length cDNA(s): AK120851"



|              |   |
|--------------|---|
| ORGANISM     | Oryza sativa (japonica cultivar-group)<br>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;<br>Ehrhartoideae; Oryzaceae; Oryza.  |
| REFERENCE    | 1   |
| AUTHORS      | Sasaki, T., Matsumoto, T. and Yamamoto, K.  |
| TITLE        | Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC<br>clone: P0415B12   |
| JOURNAL      | Published Only in Database (2002)   |
| REFERENCE    | 2   |
| AUTHORS      | Sasaki, T., Matsumoto, T. and Yamamoto, K.  |
| TITLE        | Direct Submission   |
| JOURNAL      | Submitted (20-FEB-2002) Takuji Sasaki, National Institute of<br>Agrobiological Sciences, Rice Genome Research Program; Kamondai<br>2-1-2, Tsukuba, Ibaraki 305-8602, Japan<br>E-mail: tsasaka@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,<br>Tel: 81-298-38-7441, Fax: 81-298-38-7468   |
| COMMENT      | On Jun 28, 2004 this sequence version replaced gi:37497094.<br>Genes were predicted from the integrated results of the following:<br>GENSCAN ( <a href="http://CCR-081.mit.edu/GENSCAN.html">http://CCR-081.mit.edu/GENSCAN.html</a> ), FGENESH<br>( <a href="http://www.softberry.com/">http://www.softberry.com/</a> ), GeneMark.hmm<br>( <a href="http://opal.biology.gatech.edu/GeneMark/">http://opal.biology.gatech.edu/GeneMark/</a> ), GlimmerM<br>( <a href="http://www.tigr.org/db/glimmer/glmr/form.html">http://www.tigr.org/db/glimmer/glmr/form.html</a> ), RiceHMM<br>( <a href="http://rgp.dna.affrc.go.jp/RiceHMM/">http://rgp.dna.affrc.go.jp/RiceHMM/</a> ), SplicePredictor<br>( <a href="http://bioinformatics.iastate.edu/cgi-bin/sp.cgi">http://bioinformatics.iastate.edu/cgi-bin/sp.cgi</a> ), sim4<br>( <a href="http://globin.cse.psu.edu/html/docs/sim4.html">http://globin.cse.psu.edu/html/docs/sim4.html</a> ), gap2<br>( <a href="http://www.tigr.org/software/glimmer/">http://www.tigr.org/software/glimmer/</a> ), BLASTN and BLASTX. The<br>genomic sequence was searched against NCBI NonRedundant Protein<br>database, nr ( <a href="ftp://ncbi.nlm.nih.gov/blast/db">ftp://ncbi.nlm.nih.gov/blast/db</a> ) and the cDNA<br>sequence database at RGP or DDBJ. Protein homologies of the coding<br>regions were searched against NCBI NonRedundant Protein database<br>with BLASTP. ESTs represent the identified cDNA sequences using<br>BLASTN with the corresponding DDBJ accession no. and RGP clone ID.<br>Full-length cDNAs represent the identified cDNA sequences using<br>BLASTN with the corresponding DDBJ accession no.<br>A gene with identity or significant homology to a protein is<br>classified based on the protein name to indicate the homology level<br>such as same name, 'putative-' and '-like protein'. A gene without<br>significant homology to any protein but with full-length cDNA or<br>EST homology (covering almost the entire length of partial<br>sequence) is classified as an 'unknown' protein. A gene predicted<br>by two or more gene prediction programs is classified as a<br>'hypothetical' protein according to RGPSP standard. A gene<br>predicted by a single gene prediction program is also classified as<br>a probable 'hypothetical' protein and is included as a<br>miscellaneous feature of the sequence.<br>The orientation of the sequence is from SP6 to T7 of the PAC clone.<br>This sequence of P0415B12 clone has an overlap with O3JNBa0047A17<br>clone (DDBJ: AP005749) at 5' end and an overlap with P0458B05<br>(DDBJ: AP004777) at 3' end. Detailed information on overlap and<br>assembly quality together with annotation of this entry is<br>available at <a href="http://rgp.dna.affrc.go.jp/Genomeseq.html">http://rgp.dna.affrc.go.jp/Genomeseq.html</a> .   |
| FEATURES     | Location/Qualifiers   |
| source       | 1. .157419<br>/organism="Oryza sativa (japonica cultivar-group)"<br>/mol_type="genomic DNA"<br>/cultivar="Nipponbare"<br>/db_xref="taxon:39947"<br>/chromosome="2"<br>/clone="P0415B12"<br>complement (4313. .5807)<br>/gene="P0415B12.1"<br>complement {join<4313. .4547,5568. .>5807))<br>/genes="P0415B12.1"<br>/notes="supported by full-length cDNA(s): AK121467"<br>complement (4313. .5807)<br>/genes="P0415B12.1"<br>/notes="contains full-length cDNA(s): AK121467<br>non-coding transcript<br>probably inactive due to including stop codon(s) in CDS<br>probably inactive due to too long 3'UTR in CDS"<br>6065. .7404<br>/genes="P0415B12.2"  |
| Gene         | Gene<br>complement (4313. .5807)<br>/gene="P0415B12.1"<br>complement {join<4313. .4547,5568. .>5807))<br>/genes="P0415B12.1"<br>/notes="supported by full-length cDNA(s): AK121467"<br>complement (4313. .5807)<br>/genes="P0415B12.1"<br>/notes="contains full-length cDNA(s): AK121467<br>non-coding transcript<br>probably inactive due to including stop codon(s) in CDS<br>probably inactive due to too long 3'UTR in CDS"<br>6065. .7404<br>/genes="P0415B12.2"   |
| misc_feature | misc_feature<br>complement {join<4313. .4547,5568. .>5807))<br>/genes="P0415B12.1"<br>/notes="supported by full-length cDNA(s): AK121467"<br>complement (4313. .5807)<br>/genes="P0415B12.1"<br>/notes="contains full-length cDNA(s): AK121467<br>non-coding transcript<br>probably inactive due to including stop codon(s) in CDS<br>probably inactive due to too long 3'UTR in CDS"<br>6065. .7404<br>/genes="P0415B12.2"   |
| Gene         | Gene<br>complement (4313. .5807)<br>/gene="P0415B12.1"<br>complement {join<4313. .4547,5568. .>5807))<br>/genes="P0415B12.1"<br>/notes="supported by full-length cDNA(s): AK121467"<br>complement (4313. .5807)<br>/genes="P0415B12.1"<br>/notes="contains full-length cDNA(s): AK121467<br>non-coding transcript<br>probably inactive due to including stop codon(s) in CDS<br>probably inactive due to too long 3'UTR in CDS"<br>6065. .7404<br>/genes="P0415B12.2"   |
| mrna         | mrna<br>complement {join(16514. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement (16514. .23105)<br>/note="supported by full-length cDNA(s): AK065061"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .17644,18400. .18594,<br>18674. .18751,18816. .19001,19077. .19307,19386. .19739,<br>19837. .20034,20118. .20288,20405. .20512,20611. .20829,<br>20914. .21120,21227. .21445,21564. .21638,21998. .22166,<br>22901. .23105))<br>/gene="P0415B12.4"<br>complement {join(16588. .17069,17267. .1764 |

```

/protein_id="BAD25376.1"
/db_xref="GI:49388258"
/tranlation="MASATGACGWLGRKVKVGTSGDCLLIMGSTKADVPPEKSIITLS
YLMAPLARRGVDEFPANESREFLRKLCIGKEVTPRDVDTANVGRERGTVYLGDN
VAYSIIAGARVKGPGKGGSPXYLTLLRLLEEVAKOGLGRWSKPGAAEESIRD
LPPSIIAGSFGPAKGFANVKGKSEIAIVEQVRDGSYVRVYLLPSPFQVYIVAGVQ
SPMGRRPPTVVAASADATGNGGSEEPALPTTAQRLAAAVSTEIPDRFG
IEAKHTETRLNVRRIIVVEGDTFSNIGSVYSDGDTLKDLALELVENGAKTVE
WSANMDVDAKIKLNAELQAKDOLRIWTFKPPVNTSKPIHDQKFTGKVVESGD
CIIVADAAPSGPSERRVNLSSIRAPMGNPRRDEKPNFARAEKEFLRFLISQKQ
VTVMEYSRISTGDOPTNTADAVLDYGSVFLGSPQADGDDVSSIPSSGNQPGI
NIAETLLSGFARTKRDYERESHYFDLLAAESRAEKAGVSHKSAKSPVMHITDL
TTVSAKARDFLPFLORRSHSAIVYVFSGHRFKLTIPKETSIAFSPGRCPOKD
EPYSDPALMRRIILQDVEIEVEADVDTGTFLGSLWESKTNMVASVLEAGLAKSS
FGIDRIPDANLRAEQSAKQKLTWENVYVEGESVNGSASEKOKELIKVVVTEVL
GGKFTVQTVGDRVASIQQLASLKLDKAPVIGAFNPVKGEIVLQAFSADNSWRAM
VNGPFGVSSQDDKEFVYIDYGNQVEVPYRIRPADPSISSPALAQCSLAFIKV
PNLEDFGHEAAVYLDNLSQKQYRAMIERTDTSKSGKGGTGTILITVLDAET
ETSINATMLEGRLARLSRKWDTRERKAALQNLQFOEKAKKERLQIQYGDVESDE
EQAPARRTGGR"
join(25322..25357,28498..28641)
/genes="P0415B12.5"
join(25322..25357,28498..28641)
/genes="P0415B12.5"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
29565..31528
/genes="P0415B12.6"
29565..31528
/genes="P0415B12.6"
/notes="probably inactive due to including frameshift(s) in
CDS
probably inactive due to including stop codon(s) in CDS
pseudogene, TNP2"
/pseudo
complement(31727..32496)
/genes="P0415B12.7"
complement(join(31727..32242,32434..32496))
/genes="P0415B12.7"
/notes="start and end point are not identified"
complement(join(31727..32242,32434..32496))
/genes="P0415B12.7"
/notes="similar to Oryza sativa chromosome7, P0404G11.101"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD25377.1"

Alignment Scores:
Pred. No.: 3,22e-75 Length: 157419
Score: 1148.50 Matches: 218
Percent Similarity: 84.45% Conservative: 21
Best Local Similarity: 77.03% Mismatches: 39
Query Match: 75.51% Indels: 5
DB: 8 Gaps: 3

US-10-042-894A-8 (1-289) x AP004772 (1-157419)

Qy 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLysLeu 21
Db 39928 TCCGACCTGCGCGCGGAGCACCAGGTGCGGGGCACCGCGCTCCGCGACAAAGCTG 39987

Qy 22 GlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGly 41
Db 39988 GGCCCGCTCGTGCAGCGGAGGGGCTCTTCTACAAGCCCTCCAGGCGGGGAGCGCGGG 40047

Qy 42 GluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgile 61
Db 40048 GAGCAGGAGCGCGCTTCTACCGCGCGTTCACCGCGCACCGCGCGCTCCCGCCCGCGGTC 40107

Qy 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81
Db 40108 CGGGCGCGCTTCTTCCCGCGCTTCCAGCGCACCGCGCTTCTCCCGCGCGCGCGCC 40167

```

```

Qy 82 ---GlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 40168 GCGGCGCGCGCTTACCAGCACATCGTCTCGAGGACCTCTCTCGGGGCTCCGCTCCCCC 40227

Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 40228 TGGCTCGCGAGCTCAAGATCGCGGCTGCACGTGGCGCGCGGATCCCGCGAGCCCTAC 40287

Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 40288 GTGCGCAAGTGCTTCCCAAGACCGCGAGACCAACAGCGCGCTCTCTCGGCTTCCGCGTC 40347

Qy 141 SerGlyValArgValValGlyProGlu---GlyValaValTrpArgThrGluArgProGlu 159
Db 40348 TCGCGGCTCGGGTGTTCGATCCCGGGGCGCGCTGTGGCGCGCGGACCGGTCGGAG 40407

Qy 160 ValLysAlaMetAspThrAlaGlyValaArgValLeuArgArgTyrValSerSerVal 179
Db 40408 CTGAAGGGATCGACCGCGCGGCTCCGCGGCTGCTCCGCGCTACGTTGTCACGGC 40467

Qy 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyVal 199
Db 40468 GCGCGGAGCGGCTCGGACTCGCGGCTCGCGCGCGGCTGTACGAGGAGGAGCGCGGTC 40527

Qy 200 LeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyr 219
Db 40528 CTGGCTCAGTCGGGAGCTCAAGGCTGTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 40587

Qy 220 SerAlaSerIleLeuLeuGlyTyrAsp-----AlaAlaValAlaAlaGlyGly 236
Db 40588 TCGGCTCGATTCGTTCGCTACGACGCAATCGCGCGCGCGGCTGCTCCCGGAGGT 40647

Qy 237 AspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256
Db 40648 GGAAGCGCGGCTGTAAAGGTGAAGCTGGTGGTGTTCGCGCATGTGACAGATGGGAGCGG 40707

Qy 257 ValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276
Db 40708 GTGATTGACCAACAACTTCTTGGCGGGCTCTGCTCGCTCATCAAGTTTCATCGGACATT 40767

Qy 277 ValProGlu 279
Db 40768 GTGCGCGAG 40776

RESULT 9
AK072296
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J023018G11, full
insert sequence.
ACCESSION
AK072296
VERSION
AK072296.1 GI:32982319
FLI_CDNA; CAP trapper.
KEYWORDS
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
ORGANISM
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1
AUTHORS
The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi S., Satoh K., Nagata T., Kawagashira N., Doi K.,
Kishimoto N., Yazaki J., Ishikawa M., Yamada H., Ooka H., Hotta I.,
Kojima K., Namiki T., Ohneda S., Yabagi W., Suzuki K., Li C.,
Ohtsuki K., Shishiki T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ohtsuki Y., Murakami K.,
Tada Y., Sugano S., Fujimura T., Suzuki Y., Tsunoda Y.,
Kurosaki T., Kodama T., Masuda H., Kobayashi M., Xie Q., Lu M.,
Nariawa R., Sugiyama A., Mizuno K., Yokomizo S., Niikura J.,
Ikeda R., Ishibiki J., Kawamata M., Yoshimura A., Miura J.,
Kusumegi T., Oka M., Ryu R., Ueda M., Matsubara K., Riken J.,
Kawai J., Carninci P., Adachi J., Aizawa K., Arakawa T., Fukuda S.,
Hara A., Hashidume W., Hayatsu N., Imotani K., Ishii Y., Itoh M.,

```

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
Japanese rice  
Science 301 (5631), 376-379 (2003)

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

2 (bases 1 to 1570)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
Kangawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
Masuda, H., Matubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
Numasaki, R., Oneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
Yamada, H., Yamamoto, M., Yasunishii, A., Yazaki, J., Yokomizo, S., and  
Yoshimura, A.

## TITLE

## JOURNAL

Submitted (05-DEC-2001) Shoahi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica  
rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K.,  
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and  
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory In Riken: Adachi, J., Aizawa, K.,  
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A. and Hayashizaki, Y.

## FEATURES

## source

1. .1570  
Location/Qualifiers  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023018G11"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,966-77 Length: 1570  
Score: 1144.50 Matches: 217  
Best Local Similarity: 84.10% Conservative: 21  
Query Match: 76.68% Mismatches: 40  
DB: 8 Indels: 5  
Gaps: 3  
US-10-042-894A-8 (1-289) x AK072296 (1-1570)  
QY 2 SerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeu 21  
Db 384 TCCGACCTTCGGCCCGCCGAGCACCAGGTGGGGGGGACCGCGCTCCGCCACAAGCTG 443  
QY 22 GlyProLeuIleAspGlySerClyLeuPheTyrLysProLeuGlnAlaGlyAspArgGly 41  
Db 444 GCGCCGCTCGTCGAGCGGCGAGGGGCTCTTCTACAAGCCCTCAGACCGGGGAGCGCGG 503  
QY 42 GluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIle 61  
Db 504 GAGCAGAGGGCGCTTCTACCGCGGTTTACCGCGCACCGCGCTCCGCCCGGGTCC 563  
QY 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81  
Db 564 CGGGCGGCTTCTTCCGCGCTTCCAGCGCACCGCGCTTCTCCCGCGCCCGCAGCCGC 623  
QY 82 ---GlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100  
Db 624 GCGCGCGCGCTTACCGGCACATCTCTCTCGAGACCTCTCTCGGGGCTCCCGTCCCC 683  
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTTPProProSerSerProGluProTyr 120  
Db 684 TCGGTTCGCGAGCTCAAGTTCGCGCGCTTCAGCTGGCGCGCGATCCCGGACCCCTAC 743  
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140  
Db 744 GTCCGCAAGTGCTCGCCAAAGGACCGAGACCAAGCGCGCTCTCTCGGCTTCCGCGTC 803  
QY 141 SerGlyValArgValValGlyProGlu---GlyAlaValTTPArgThrGluArgProGlu 159  
Db 804 TCCGCGCTCGCGGTGGTTCGATCCCGGGGCGCGCGCTGTGGCGCCCGGACCGGTTCGAG 863  
QY 160 ValLysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerVal 179  
Db 864 CTGAGAGGATCGACCGCGCGGGTCCGCGCGTGTCTCCGCGCTACGTGTCCACGGGC 923  
QY 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyClyLysGlyVal 199  
Db 924 GCGCGGAGCGGCTGGACTCGCGCTCGCGCGCGGTGTACGAGGGGAGGCGCGCGTC 983  
QY 200 LeuSerGlnLeuArgGluLeuLysAlaTTPPheGluGluGlnThrLeuPheHisPheTyr 219  
Db 984 CTGGTTCAGTTCGGGAGCTCAAGCGGTGTTCGAGGAGCAAAACCTCTTACCATTCTAC 1043  
QY 220 SerAlaSerIleLeuLeuGlyTyrAsp-----AlaAlaAlaValAlaAlaGlyGly 236  
Db 1044 TCGGGTTCGATTCTGTTCGGCTACGACCGCATTCGGCGCGCGCGCTGTCTCCCGAGGT 1103  
QY 237 AspGlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256  
Db 1104 GGAAGCGCGGTGTGAAGGTGAAGTGTGTGTGAGTTCGCGCATGTTCGACAGGGGAGCGG 1163  
QY 257 ValIleAspHisAsnPheLeuGlyCysSerLeuIleLysPheValSerAspIle 276  
Db 1164 GTGATTGACCAACTTCTTTGGGCGGCTCTCTCTGCTCATCAAGTTTCATCGCGCAATT 1223  
QY 277 ValProGlu 279  
Db 1224 GTCCGCGAG 1232  
AXS13580 643 bp DNA linear PAT 05-OCT-2002  
RESULT 10  
AXS13580  
LOCUS

```

DEFINITION   Sequence 17 from Patent WO02059324.
ACCESSION    AX513580
VERSION      AX513580.1  GI:23559681
KEYWORDS
SOURCE       Zea mays
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1
AUTHORS      Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE        Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL      PIONEER HI-BRED INTERNATIONAL, INC. (US)
              Patent: WO 02059324-A 17 01-AUG-2002;
              PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES     source
              1. .643
              /organism="Zea mays"
              /mol_type="unassigned DNA"
              /db_xref="taxon:4577"
ORIGIN
Alignment Scores:
Pred. No.:      2,45e-52      Length:      643
Score:          808.50      Matches:    162
Percent Similarity: 93.14%      Conservative: 1
Best Local Similarity: 92.57%      Mismatches: 10
Query Match:    53.16%      Indels:     3
DB:             Gaps:        0

US-10-042-894A-8 (1-289) x AX513580 (1-643)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisLys 20
DB 117 ATCCCGACGCTCCACCGCGCGAGCACCAAGTCGCGCGTCCACCGCGCTCCGCGCAAG 176
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArg 40
DB 177 CTGGGCGCGCTCATCGACGCTCGCGCTCTTCTACAGCGCTCCAGCGCGCGCGCGT 236
QY 41 GlyGluHisGluValAlaPheTyrlsGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 237 GGGGAGCAGCGGTCGCTCTATGAGCGCTTCTCGCGCCACCGCGCGTCCGCGCGCG 296
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 297 ATCCGAGACACCTTCTTCCCGCGGTTCCACGGCAGCGACTCTCCCGCACCGGCGCAG 356
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 357 CCCGGGAGCGGATCCGACCTCTCTCGACGCTCTCGCGCGGTTTGAAGCGCCC 416
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB 417 TGGCTCGCAGACATCAAGATCGCGCCATCACGTGGCCACCGAGTTCGCGGAGCCCTAC 476
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 477 ATCCNCAAGTACCTNGCCCAAGGACCGCGGACCAAGCGTCTCTGCTCGGATTCGCGTC 536
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArg-ProGluVa 160
DB 537 TTGC--GTCCGAGTGTGCGCGCGGAGGCGCGTGTGCGGACGAGGCGCGCGGGT 594
QY 160 lLysAlaMetAspThr-AlaGlyValArgValLeuArg 173
DB 595 GAANGCTATGACACCGCTCGGNGNCCGGGNGTGTCTCGG 635

RESULT 11
LOCUS       AX513576
DEFINITION Sequence 13 from Patent WO02059324.
ACCESSION   AX513576
VERSION     AX513576.1  GI:23559677
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .643
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Alignment Scores:
Pred. No.:      2,45e-52      Length:      643
Score:          808.50      Matches:    162
Percent Similarity: 93.14%      Conservative: 1
Best Local Similarity: 92.57%      Mismatches: 10
Query Match:    53.16%      Indels:     3
DB:             Gaps:        0

US-10-042-894A-8 (1-289) x AX513576 (1-1020)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerHisLysLeuGlyPro 23
DB 24 CTCAGGCGCCAGATCATCAGGTTCGTGGACATGAAGCTGGGCTCGGGAAGCTTGGCCCA 83
QY 24 LeuIleAspGlySerGlyLeuPheTyrlsProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 84 CTCATTGATGATTACGCGCGGTTTACAAACCACTGACGGTGATAACCGTGGGTAGAA 143
QY 44 GluValAlaPheTyrlsGluAlaPheSerAlaHisAlaAlaValProAlaAspGlyArg 63
DB 144 GAAGTAGCCTTTTATGAATCATTTTCTTACAAATAATATTCAGNACACATACGC--- 200
QY 64 ThrPheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
DB 201 AAATCTTCTTATATATTATGGCACCACAAATCATG-----AAGGCATCCACTGGCTCT 254
QY 84 ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
DB 255 GACCATCTCCATCATGTTTGAAGATCTTACATCATGCTCATGTCAACCCATCTGTAATG 314
QY 104 AspIleValIleGlyAlaIleThrTrpProSerSerProGluProTyrIleAlaLys 123
DB 315 GACATCAAAATCGGGTCCAGAACATGGCGCCAGAAAGCTTCCAGGGGTACATTGCAAAA 374
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
DB 375 TGGTTAAAAAAGGATAGGGAACACACAGATATTCCATTGGGATTCAGGATCTCGGCGCTG 434
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
DB 435 CAACTCTATATCATGATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 494
QY 163 MetAspThrAlaGlyValArgArgValLeuArgValLeuArgValSerValAlaAspGlu 182

```















GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 21:48:51 ; Search time 5916 Seconds  
(without alignments)  
11008.082 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgagtcgctgcac.....ataaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sus.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 1344  | 100.0       | 1344   | 6     | AX513570 Sequence |
| 2          | 865.4 | 64.4        | 923    | 6     | AX513568 Sequence |
| 3          | 844.8 | 62.9        | 1169   | 6     | AX513564 Sequence |
| 4          | 837.4 | 62.3        | 923    | 6     | AX513566 Sequence |
| 5          | 740.4 | 55.1        | 3416   | 6     | AX513583 Sequence |
| 6          | 686   | 51.0        | 899    | 6     | AX513578 Sequence |
| 7          | 546.4 | 40.7        | 149142 | 8     | AP005749          |
| 8          | 546.4 | 40.7        | 157419 | 8     | AP004772          |
| 9          | 544.8 | 40.5        | 1570   | 8     | AX072296          |
| 10         | 470.4 | 35.0        | 643    | 6     | AX513580          |
| 11         | 298.4 | 22.2        | 519    | 6     | AX513581          |
| 12         | 191.6 | 14.3        | 353    | 6     | AX513582          |
| 13         | 184.6 | 13.7        | 1195   | 6     | AX513574          |
| 14         | 181.6 | 13.5        | 25054  | 8     | LUS310150         |
| 15         | 179   | 13.3        | 1105   | 6     | AX513572          |
| 16         | 149.8 | 11.1        | 1020   | 6     | AX513576          |
| 17         | 148.6 | 11.1        | 1390   | 8     | AY136378          |
| 18         | 148.6 | 11.1        | 116763 | 8     | ATT211            |
| 19         | 148.2 | 11.0        | 905    | 8     | BT000196          |

|    |       |      |        |    |              |
|----|-------|------|--------|----|--------------|
| 20 | 147.8 | 11.0 | 115040 | 8  | AC124954     |
| 21 | 146.6 | 10.9 | 858    | 8  | BT011900     |
| 22 | 146.6 | 10.9 | 1243   | 8  | AY087217     |
| 23 | 145.8 | 10.8 | 1068   | 8  | ATH404678    |
| 24 | 145.8 | 10.8 | 130078 | 2  | AC127018     |
| 25 | 145.6 | 10.8 | 903    | 8  | ATH243592    |
| 26 | 145.6 | 10.8 | 903    | 8  | AY072621     |
| 27 | 145.6 | 10.8 | 903    | 8  | AY147936     |
| 28 | 145.6 | 10.8 | 1166   | 8  | AF412073     |
| 29 | 145.6 | 10.8 | 1170   | 8  | ATH245521    |
| 30 | 145.6 | 10.8 | 57246  | 8  | AB010069     |
| 31 | 145.4 | 10.8 | 861    | 8  | AY147935     |
| 32 | 143.6 | 10.7 | 1130   | 8  | AY085862     |
| 33 | 137.6 | 10.2 | 84869  | 12 | AY350714     |
| 34 | 83.8  | 6.2  | 43428  | 8  | AP004981     |
| 35 | 75.2  | 5.6  | 110000 | 2  | LMFLCHR32_06 |
| 36 | 75    | 5.6  | 110000 | 2  | LMFLCHR36_03 |
| 37 | 74.8  | 5.6  | 296500 | 1  | SCO939128    |
| 38 | 73.8  | 5.5  | 125020 | 9  | AF429315     |
| 39 | 73.6  | 5.5  | 110000 | 2  | LMFLCHR36_31 |
| 40 | 73.4  | 5.5  | 300800 | 1  | SCO939112    |
| 41 | 72.4  | 5.4  | 113193 | 1  | AF357202     |
| 42 | 72.4  | 5.4  | 113193 | 6  | AX703543     |
| 43 | 71.4  | 5.3  | 135638 | 1  | AF484556     |
| 44 | 70.4  | 5.2  | 224    | 8  | AY022706     |
| 45 | 70.4  | 5.2  | 78220  | 2  | AC023212     |

#### ALIGNMENTS

|            |            |                                    |         |     |        |                 |
|------------|------------|------------------------------------|---------|-----|--------|-----------------|
| RESULT 1   | AX513570   | Sequence 7 from Patent WO02059324. | 1344 bp | DNA | linear | PAT 05-OCT-2002 |
| LOCUS      | AX513570   |                                    |         |     |        |                 |
| DEFINITION | AX513570   |                                    |         |     |        |                 |
| ACCESSION  | AX513570   |                                    |         |     |        |                 |
| VERSION    | AX513570.1 | GI:23559670                        |         |     |        |                 |
| KEYWORDS   |            |                                    |         |     |        |                 |
| SOURCE     |            |                                    |         |     |        |                 |
| ORGANISM   |            |                                    |         |     |        |                 |

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1  
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.  
TITLE Novel inositol polyphosphate kinase genes and uses thereof  
JOURNAL Patent: WO 02059324-A 7 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES  
source Location/Qualifiers  
1..1344  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
52..921  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52954.1"  
/db\_xref="GI:23559671"  
/translation="MSDLHPPEHQVAGHRASAKGLPLIDSGSLFYKPLQADGRGEHE  
VAFYFAFSAHAAPARIDTFFPHGHTRLLPTEAQGPGEHPHVLDDLLAGFOAPCV  
ADIKIGAITWPPSPPEPIYAKLAKDRGTTSVLLGFRVGVVVEGAVMTERPEV  
KAMPTAGVRVLRVSVSVADEGMDCAAAYVGGGVLSQLRELKAWFEQTLFHF  
YMSILIGYDAAVAAGDGGVTVKLVDFAHVAEGDGVIDHNFGLGSLIKFVSDI  
VPETPTQPLGPS"

#### CDS

#### ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 6.8e-201;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCACGAGTCCGTCGTCACCCCTCGGCCCATAGTCCCTCCCATACCATGTCGCAC 60  
Db 1 GCACGAGTCCGTCGTCACCCCTCGGCCCATAGTCCCTCCCATACCATGTCGCAC 60

```
Qy 61 CTCACCGCGGAGCACCAGTCCCGCGCACCGCGCTCCCGCAGCAGCTGGGCGG 120
Db 61 CTCACCGCGGAGCACCAGTCCCGCGCACCGCGCTCCCGCAGCAGCTGGGCGG 120
Qy 121 CTATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGCGCGGAGCCGCTGGGGAGCAC 180
Db 121 CTATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGGCGCGGAGCCGCTGGGGAGCAC 180
Qy 181 GAGTCGCTCTATAGAGGGTTCTCCGCCCAACCGCGCGTCCCGCGCGCATCCGAGAC 240
Db 181 GAGTCGCTCTATAGAGGGTTCTCCGCCCAACCGCGCGTCCCGCGCGCATCCGAGAC 240
Qy 241 ACCTTCTTCCCGGTTCCAGGACCGGACTCTCCCAACCGAGGCGGAGCCCGGGAG 300
Db 241 ACCTTCTTCCCGGTTCCAGGACCGGACTCTCTCCCAACCGAGGCGGAGCCCGGGAG 300
Qy 301 CCGCATCTCACCCTCGCTCCAGACCTCTCCGGGGTTTCAGGGCGCTCGGTCGCA 360
Db 301 CCGCATCTCACCCTCGCTCCAGACCTCTCCGGGGTTTCAGGGCGCTCGGTCGCA 360
Qy 361 GACATCAAGATCGCGGCATCAGTGGCCACCGAGTTGCGCGGAGCCCTACATCGCCAAG 420
Db 361 GACATCAAGATCGCGGCATCAGTGGCCACCGAGTTGCGCGGAGCCCTACATCGCCAAG 420
Qy 421 TGCCTCGCAGGACCGGGACCAAGCGGTTCTCGGATTCGGGCTCCGGCGTC 480
Db 421 TGCCTCGCAGGACCGGGACCAAGCGGTTCTCGGATTCGGGCTCCGGCGTC 480
Qy 481 CGAGTCGTCCGCCCGGAGGCGCGTGTGGCGGACCGAGCGCCCGGAGGTGAAGGCCATG 540
Db 481 CGAGTCGTCCGCCCGGAGGCGCGTGTGGCGGACCGAGCGCCCGGAGGTGAAGGCCATG 540
Qy 541 GACACGCGCGGCTCCCGCGTGTCCGGCGTACGTTGTCATCGTTGCGGACGAGGG 600
Db 541 GACACGCGCGGCTCCCGCGTGTCCGGCGTACGTTGTCATCGTTGCGGACGAGGG 600
Qy 601 ATGAGCTGTGGCTCCCGCGCGCGGTGTACGAGGAAAGTGGAGTCTGTACACGTG 660
Db 601 ATGAGCTGTGGCTCCCGCGCGCGGTGTACGAGGAAAGTGGAGTCTGTGTACACGTG 660
Qy 661 CGGAGCTCAAGGCGTGTTCAGGAGCAGACTCTGTTCCACTTCTACTCGGCGTCGATT 720
Db 661 CGGAGCTCAAGGCGTGTTCAGGAGCAGACTCTGTTCCACTTCTACTCGGCGTCGATT 720
Qy 721 CTTCTGGGCTATGATGCTGTCAGTCCAGACGCGGAGATGGGGTGGGGTGAACGGTG 780
Db 721 CTTCTGGGCTATGATGCTGTCAGTCCAGACGCGGAGATGGGGTGGGGTGAACGGTG 780
Qy 781 AAGCTGTGGACTTTGCCCATGTGCCGAGGGTGTGGGGTGTATGACCAACTTCTGT 840
Db 781 AAGCTGTGGACTTTGCCCATGTGCCGAGGGTGTATGGGGTGTATGACCAACTTCTGT 840
Qy 841 GCGGGCTCTCTCGCTGATCAAGTTCGTTCTGACATGTTCCGGAGACTCTCTATACG 900
Db 841 GCGGGCTCTCTCGCTGATCAAGTTCGTTCTGACATGTTCCGGAGACTCTCTATACG 900
Qy 901 CAGCCTTTGGGTCTCTTAAGAGAGGATCCTGGCATTTGATTTGATAACAAAGCCCTA 960
Db 901 CAGCCTTTGGGTCTCTTAAGAGAGGATCCTGGCATTTGATTTGATAACAAAGCCCTA 960
Qy 961 CAAGTTTGTCTGAAAAGAGCGCTCCAGTTGTGCTGGGTGTGGAGATCTGAGACG 1020
Db 961 CAAGTTTGTCTGAAAAGAGCGCTCCAGTTGTGCTGGGTGTGGAGATCTGAGACG 1020
Qy 1021 GTGTCGSCCACTTGGTGTGCTTTCCTTTCCTTTCCTGCAAAATACAGCAACCTGCT 1080
Db 1021 GTGTCGSCCACTTGGTGTGCTTTCCTTTCCTTTCCTGCAAAATACAGCAACCTGCT 1080
Qy 1081 CTTTTTTTCGCAACCCCTTACTTCGAGAGAACTTTTTTTTTTTCCTTTCCTTTCGAG 1140
Db 1081 CTTTTTTTCGCAACCCCTTACTTCGAGAGAACTTTTTTTTTTTCCTTTCCTTTCGAG 1140
```

```
Qy 1141 TTACGTTGGATCTGGTTTGTCCACTCGGTTCAGAGGTTGTAAGCATGAGGAGCGGTG 1200
Db 1141 TTACGTTGGATCTGGTTTGTCCACTCGGTTCAGAGGTTGTAAGCATGAGGAGCGGTG 1200
Qy 1201 TGATCCGGCAACTGTGTGTCAGTCTTTGCGTGTCTGCGGTTTTCGATGGCTTTGCTGC 1260
Db 1201 TGATCCGGCAACTGTGTGTCAGTCTTTGCGTGTCTGCGGTTTTCGATGGCTTTGCTGC 1260
Qy 1261 TCGATCCGATGTGTACTGAGATCGTAGTGAGGAGTCTTCTACCTCCAAAGCAATCCG 1320
Db 1261 TCGATCCGATGTGTACTGAGATCGTAGTGAGGAGTCTTCTACCTCCAAAGCAATCCG 1320
Qy 1321 TCCGATAAAAAAATAAAAAA 1344
Db 1321 TCCGATAAAAAAATAAAAAA 1344

RESULT 2
AX513568
LOCUS AX513568
DEFINITION Sequence 5 from Patent WO02059324.
ACCESSION AX513568
VERSION AX513568.1
KEYWORDS GI:23559668
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 5 01-AUG-2002, (US)
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..922
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD52953.1"
/db_xref="GI:23559669"
/translation="MEDLPPEHQVAGHRASAKPGPLIDGSLFYKPLQAGDGEHE
VAFYEAFSAHAAPARIDTFFRPHGRLLEAQPGEHPHLVLDLLAGAEPCV
ADIKGATWPPSSPEFYAKLMDRGTTSVLLGFRVSGRVVGPGEVWRTREPV
KAMDTAGVRRVLRYSVADEGMDCALAAVYGGKGVLSQLRELKAWFEQTLFPH
YSASILGYDAAAFAAGGGGGVTKLVDPFHVAEAGDVIDHNFGLGSLIKFVSDI
VPETPQTQPLGPS"
ORIGIN
Query Match 64.4%; Score 865.4; DB 6; Length 923;
Best Local Similarity 98.8%; Pred. No. 6.1e-126;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 42 TCCCATACCATGTCGACCTCCACCGCGGAGCACAAGTCGCGGCGCACCGCGCTC 101
Db 43 TCCCATACCATGTCGACCTCCACCGCGGAGCACAAGTCGCGGCGCTC 102
Qy 102 CGCAGCAAGCTGGGCGGCTCATCGACGCTCGGCTCTTCTACAAGCGCTCCAGGC 161
Db 103 CGCAGCAAGCGGGCGGCTCATCGACGCTCGGCTCTTCTACAAGCGCTCCAGGC 162
Qy 162 CGGCGACCGTGGGAGCACGAGTTCGCTTTCTATGAGCGTTCCTCCGCCACGCGCGT 221
Db 163 CGGCGACCGTGGGAGCACGAGTTCGCTTTCTATGAGCGTTCCTCCGCCACGCGCGT 222
Qy 222 CGGCGCGGATCGGAGACCTTCTCCCGCGGTTCACGCGACGAGTCTCTCCCCAC 281
Db 223 CGGCGCGGATCGGAGACCTTCTTCCCGCGGTTCACGCGACGAGTCTCTCCCCAC 282
Qy 282 CGAGGCGCAGCGCGGAGCGGATCTCTACCTCTCGACGACCTCTCTCGGGGTT 341
```

```
Db 283 CGAGCGCAGCCCGGGAGCGGCATCCGACCTCGTCTCGACGACCTCTCGGGGATT 342
Qy 342 TCAGCGCCCTGCGTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGATTGCGC 401
Db 343 TGAGCGCCCTGCGTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGATTGCGC 402
Qy 402 GGAGCCCTTACATCGCAAGTGCCTCGCCCAAGGACCGCGGACCCAGAGCGTTCTGCTCGG 461
Db 403 GGAGCCCTTACATCGCAAGTGCCTCGCCCAAGGACCGCGGACCCAGAGCGTTCTGCTCGG 462
Qy 462 ATTCCGCTCTCCGCGCTCGAGTGCCTCGGCGCCGAGGCGCGTGTGGCGGACCGAGCG 521
Db 463 ATTCCGCTCTCCGCGCTCGAGTGCCTCGGCGCCGAGGCGCGTGTGGCGGACCGAGCG 522
Qy 522 CCCGAGGTGAAGGCATGGAACACCGCGCGGTCCGCGCGTGTCCGCGCGTGTGCTGTC 581
Db 523 CCCGAGGTGAAGGCATGGAACACCGCGCGGTCCGCGCGTGTCCGCGCGTGTGCTGTC 582
Qy 582 ATCCGTTCCCGACGAGGGGATGGAATGTCGCTCGCGCGCGGTGTACGGAGGAAAAG 641
Db 583 ATCCGTTCCCGACGAGGGGATGGAATGTCGCTCGCGCGCGGTGTACGGAGGAAAAG 642
Qy 642 TGGAGTCTTGTACAGTGTGCGGAGCTCAAGGCGTGTTCGAGGAGAGACTCTGTTCCA 701
Db 643 TGGAGTCTTGTACAGTGTGCGGAGCTCAAGGCGTGTTCGAGGAGAGACTCTGTTCCA 702
Qy 702 CTCTACTCGCGCTCGATTCTCTGGGCTATGATGCTGCTCGAGTCGCGAGCGGAGA 761
Db 703 CTCTACTCGCGCTCGATTCTCTGGGCTATGATGCTGCTCGAGTCGCGAGCGGAGG 762
Qy 762 TGGGGTGGGTGACGCTGAAGCTGCTGGAATTTCCCATGTTGCGCGAGGGTGTGCGGT 821
Db 763 TGGGGTGGGTGACGCTGAAGCTGCTGGAATTTCCCATGTTGCGCGAGGGTGTGCGGT 822
Qy 822 GATTGACCACAACTTCTCGGCGGCTCTGCTCGCTGATCAAGTTTCTGACATTGT 881
Db 823 GATTGACCACAACTTCTCGGCGGCTCTGCTCGCTGATCAAGTTTCTGACATTGT 882
Qy 882 TCCGAGACTCTCTATAGCGAGCTTTGGGTCTCTTTAAG 922
Db 883 TCCGAGACTCTCTAGAGCGAGCTTTGGGTCTCTTTAAG 923
```

RESULT 3  
AX513564  
LOCUS 1169 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 1 from Patent WO2059324.  
ACCESSION AX513564  
VERSION AX513564.1 GI:23559664

KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1  
Shi, J., Beach, L. R., Wang, H., Rafalski, J. A. and Cahoon, R. E.  
Novel inositol polyphosphate kinase genes and uses thereof.  
Patent: WO 02059324-A 1 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES  
Location/Qualifiers  
1..1169  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
84..806  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52951.1"  
/db\_xref="GI:23559665"  
/translation="MPDLHPPEHQVAGHRASKLGLPLIDGSLFYKPLQAGDRGEHE  
VAFYEAFAHAAPVARIIRDTPFRFHGTRLLPTEAPGEPHPHLVLDLLAGFEAPCV

ADIKIGAITWPPSPPEYIAKYLAQRGTTSVLLGFRVLRPRRPRRRVADGAPGGE  
GYHRRFPALRVIRLPTGRTABRRRRTBEKVESHSCASSRHGWRSLCSTS  
TRRRFFWAMMLLOSQAQEVGVG"  
ORIGIN  
Query Match 62.9%; Score 844.8; DB 6; Length 1169;  
Best Local Similarity 97.2%; Pred. No. 1e-122;  
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;  
Qy 42 TCCCATACCATGTCGACCTCCACCGCGCGGAGCACCAAGTCGCGGCGCACCGCGCTC 101  
Db 74 TCCCATACCATGTCGCGGACCTCCACCGCGCGGAGCACCAAGTCGCGGCTCACCGGCTC 133  
Qy 102 CGCCAGCAAGCTGGGCGCGCTCATCGACGGCTCGGCGCTCTTCTAAGCGCTCCAGGC 161  
Db 134 CGCCAGCAAGCTGGGCGCGCTCATCGACGGCTCGGCGCTCTTCTAAGCGCTCCAGGC 193  
Qy 162 CGCGGAGTGGGAGACGAGGTGCGCTTCTATGAGGCGTCTTCGCGCGGCGCGCT 221  
Db 194 CGCGGAGTGGGAGACGAGGTGCGCTTCTATGAGGCGTCTTCGCGCGGCGCGCT 253  
Qy 222 CGCGGCGCATCGAGACACTTCTTCCCGCGGTTCACGCGGCGCGGCTCTCTCCCGAC 281  
Db 254 CGCGGCGCATCGAGACACTTCTTCCCGCGGTTCACGCGGCGCGGCTCTCTCCCGAC 313  
Qy 282 CGAGCGCGAGCGCGGAGCGCATCTCTACTCTCTCTCGACGACTCTCTCGCGGCTT 341  
Db 314 CGAGCGCGAGCGCGGAGCGCGCATCTCTACTCTCTCTCGACGACTCTCTCGCGGCTT 373  
Qy 342 TCAGCGCGCTGCTGCGAGACATCAAGATCGGCGCGCATCAAGTGGCGCGGCTCTCGC 401  
Db 374 TGAGCGCGCTGCTGCGAGACATCAAGATCGGCGCGCATCAAGTGGCGCGGCTCTCGC 433  
Qy 402 GGAGCGCTTACATCGCAAGTACCTCGCGGCGCGGAGCGCGGCTCTCTCGG 461  
Db 434 GGAGCGCTTACATCGCAAGTACCTCGCGGCGCGGAGCGCGGCTCTCTCGG 493  
Qy 462 ATTCCGCTCTCTCGCGCTCGAGTCTCGCGCGCGGCGCGGCTCTCTCGGCGAGCG 521  
Db 494 ATTCCGCTCTCTCGCGCTCGAGTCTCGCGCGCGGCGCGGCTCTCTCGGCGAGCG 551  
Qy 522 CCCGAGGTGAAGCGCATGGAACACCGCGCGCGCTCGCGCGCGGCTCTCTCGGCGAG 581  
Db 552 CCCGAGGTGAAGCGCATGGAACACCGCGCGCGCTCGCGCGCGGCTCTCTCGGCGAG 611  
Qy 582 ATCCG-TTGGCGAGGAGGATGAGTGTGCGCTCGCGCGCGGCTGTACGAGGAGAAAG 640  
Db 612 ATCCGCTTGGCGAGGAGGATGAGTGTGCGCGCTCGCGCGCGGCTGTACGAGGAGAAAG 671  
Qy 641 GTGAGTCTTGTACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGAGGAGTCTGTTC 700  
Db 672 GTGAGTCTTGTACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGAGGAGTCTGTTC 731  
Qy 701 ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGTGCTGAGTGTGCGAGCGGAG 760  
Db 732 ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGTGCTGAGTGTGCGAGCGGAG 791  
Qy 761 ATGGGGTGGGTCAGCGTGAAGTGTGGGCTTTGCCCATGTGGCGGCTGTATGGG 820  
Db 792 GTGGGGTGGGTCAGCGTGAAGTGTGGGCTTTGCCCATGTGGCGGCTGTATGGG 851  
Qy 821 TGATTGACCACAACTTCTCGCGCGGCTCTGCTGCTGATCAAGTGTCTGTGACATTG 880  
Db 852 TGATTGACCACAACTTCTCGCGGAGCTCTGCTAGTGTCAAGTGTCTGTGACATTG 911  
Qy 881 TTCCGAGACTCTCTATACGAGCGCTTTGGGTCTCTTAAAGAGAGGATCTCTGGCA-TTT 939  
Db 912 TTCCAGAGACTCTTTAGAGCGAGCTTTGGGTCTCTTAAAGAGAGGATCTCTGACATTTT 971  
Qy 940 CGATTGTGATAAAG 955  
Db 972 TGATTGATAAAG 987



```
RESULT 4
AX513566
LOCUS AX513566 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 3 from Patent WO02059324.
ACCESSION AX513566
VERSION AX513566.1 GI:23559666
KEYWORDS
SOURCE
ORGANISM Zea mays
Ze mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 3 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..736
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52952.1"
/db_xref="GI:23559667"
/translation="MPDLHPPEHQVAGHRASAKPGPLIDSGLYFKPLQAGDRGEHE
VAFYEAFSAHAAPARIDTFPPFHGTRLLPTPEAQEPHPLVLDDLLAGFEAPCV
ADIKIGALTWPPSSPEPIYAKCLAMDRGTTSVLLGPRVSGVRVVPGEAVRTPEPV
KAMDTGVRRVLRVYSSACRRGDLRARGGVRRKXSLVTAARAQGVVRGAASVPL
LLGVDDSSGL"
ORIGIN
Query Match 62.3%; Score 837.4; DB 6; Length 923;
Best Local Similarity 97.5%; Pred. No. 1.5e-121;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
Qy 42 TCCCATACCATGTCGAGACACCTTTCTTCCCGGTTCCACGGACGCGACTCTCTCCCGAC 101
Db 43 TCCCATACCATGTCGAGACACCTTTCTTCCCGGTTCCACGGACGCGACTCTCTCCCGAC 102
Qy 102 CGCAGCAGTGGGCGGCTCATCGAGCGGTCGGGCTCTTCTACAGCGGCTCCAGGC 161
Db 103 CGCAGCAGTGGGCGGCTCATCGAGCGGTCGGGCTCTTCTACAGCGGCTCCAGGC 162
Qy 162 CGGCGACCGTGGGAGACAGAGTTCGCTTCTATGAGCGGTTCTCCGCCACGCGCGGT 221
Db 163 CGGCGACCGTGGGAGACAGAGTTCGCTTCTATGAGCGGTTCTCCGCCACGCGCGGT 222
Qy 222 CCGCGCGCATCGGACACCTTTCTTCCCGGTTCCACGGACGCGACTCTCTCCCGAC 281
Db 223 CCGCGCGCATCGGACACCTTTCTTCCCGGTTCCACGGACGCGACTCTCTCCCGAC 282
Qy 282 CGAGCGGACCGCGGAGCGCGATCTCTCACTGCTCTCGACGACCTCTCGCGGGTT 341
Db 283 CGAGCGGACCGCGGAGCGCGATCTCGACCTCTCTCTCGACGACCTCTCTCGCGGATT 342
Qy 342 TCAGCGCGCTGCGTTCGAGACATCAAGATCGGCGCATCATCGTGGCCACCGAGTTCGCC 401
Db 343 TGAGCGCGCTGCGTTCGAGACATCAAGATCGGCGCATCATCGTGGCCACCGAGTTCGCC 402
Qy 402 GGAGCGCTTACATCGCCAGTGCCTCGCAAGAGACCGGGGACCAAGAGCGTTCTGCTCGG 461
Db 403 GGAGCGCTTACATCGCCAGTGCCTCGCCATCGACCGGGGACCAAGAGCGTTCTGCTCGG 462
Qy 462 ATTGGGCTTCCGGGCTCCAGTTCGTCGGCCCGAGGGCCCGTGTGGCGGACGAGCG 521
Db 463 ATTGGGCTTCCGGGCTCCAGTTCGTCGGCCCGAGGGCCCGTGTGGCGGACGAGCG 522
Qy 522 CCGGAGGTGAAGGCATGACACCGCGCGGTCGCGCGGCTCTTCTACAGCGGCTCCAGCGGAGCAGC 581
```

```
Db 523 CCGGAGGTGAAGGCTATGGACACCGTGGCGTCCGCCGCTGCTCCGGCGCTACGTGTC 582
Qy 582 ATCCG-TTGCAGAGGAGATGAGACTGCGGCTCGCGCGCGCGGCTTACGAGGAAAG 640
Db 583 ATCCGCTTGCAGAGGAGATGAGACTGCGGCTCGCGCGCGCGGCTTACGAGGAAAG 642
Qy 641 GTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGCTGTTTCGAGGAGCAGACTCTGTTC 700
Db 643 GTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGCTGTTTCGAGGAGCAGCTCTGTTC 702
Qy 701 ACTTACTCGCGGTCGATTTCTTCTGGGCTATGATGCTGTCAGTCGACGAGCGGAG 760
Db 703 ACTTACTCGCGGTCGATTTCTTCTGGGCTATGATGCTGTCAGTCGACGAGCGGAG 762
Qy 761 ATGGGGTGGGTCAGCGTGAAGCTGGTGAAGCTTTGCCATGTGGCCGAGGTCATCGGG 820
Db 763 GTGGGGTGGGTAACAGTGAAGCTGGTGAAGCTTTGCCATGTGGCCGAGGTCATCGGG 822
Qy 821 TGATTGACCAACAATTCTCTGGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTG 880
Db 823 TGATTGACCAACAATTCTCTGGCGGCTCTGCTAGTGAATCAAGTTCGTTTCTGACATTG 882
Qy 881 TTCGGAGACTCTCATACGACGCTTTGGGTCCTTCTTAA 921
Db 883 TTCAGAGACTCTCTCAGACGCGCTTTGGGTCCTTCTTAA 923
RESULT 5
AX513583
LOCUS AX513583 3416 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 20 from Patent WO02059324.
ACCESSION AX513583
VERSION AX513583.1 GI:23559684
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 20 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..3416
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
72..407
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52959.1"
/db_xref="GI:23559685"
/translation="MPDLHPPEHQVAGHRASAKPLIDSGLYFKPLQAGDRGEHE
VAFYEAFSAHAAPARIDTFPPFHGTRLLPTPEAQEPHPLVLDDLLAGFEAPCV
ADIKIGAIT"
ORIGIN
Query Match 55.1%; Score 740.4; DB 6; Length 3416;
Best Local Similarity 90.2%; Pred. No. 2e-106;
Matches 862; Conservative 0; Mismatches 31; Indels 53; Gaps 4;
Qy 2 CACGAGTCAGTCCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGCACC 61
Db 28 CACCATCGCACCGCTCACCCCTTGCTCCCATAG-----TCCCATACCATGTCGCACC 81
Qy 62 TCACCGCGCGGAGCACCAAGTCGCGCCACCGCGCTCGCCAGCAAGCTGGGCGCGC 121
Db 82 TCACCGCGCGGAGCACCAAGTCGCGCGCTCACCGCGCTCGCCAGCAAGCTGGGCGCAC 141
Qy 122 TCATCGAGCGTCCGGCTCTTCTACAGCGGCTCCAGGCGCGCACCGTGGGAGCAGC 181
```

Db 142 TCATCGACGACTGCGCTCTTCTTACAAGCGCTCCAGCGCGGACCGTGGGGAGCAGC 201  
Qy 182 AGGTGCGCTTCTATGAGCGGTTCCTCGCCACGCGCGCTCCGCGCCGCGCATCCGAGACA 241  
Db 202 AGGTGCGCTTCTATGAGCGGTTCCTCGCCACGCGCGCTCCGCGCCGCGCATCCGAGACA 261  
Qy 242 CTTCTTCTCCCGGTTCCACGCGACGCTCTCTCCACCGAGCGCGAGCCCGGGAGC 301  
Db 262 CTTCTTCTCCCGGTTCCACGCGACGCTCTCTCCACCGAGCGCGAGCCCGGGAGC 321  
Qy 302 CGCATCTCTACCTGCTCTCGAGACCTCTCTCGCGGGTTTCAGCGCGCTCTGCTCGAG 361  
Db 322 CGCATCTCTACCTGCTCTCGAGACCTCTCTCGCGGGTTTCAGCGCGCTCTGCTCGAG 381  
Qy 362 ACATCAAGATCGGCGCATACGCTGCGCCACCGAGTTGCGCGGAGCCCTACATCGCAAGT 421  
Db 382 ACATCAAGATCGTGCATACAGT----- 405  
Qy 422 GCCTCGCCAAAGACCGCGGACACAGAGGTTCTGCTCGGATTCGCGCTCTCCGCGTCC 481  
Db 406 -----GACCACGAGGTTCTGCTCGGATTCGCGTCTCCGCGTCC 446  
Qy 482 GAGTCGTGCGGCCCGAGCGCGCTGTGCGGACGCGAGCGCGCGAGGTGAAGGCCATGG 541  
Db 447 GAGTCGTGCGGCCCGAGCGCGCTGTGCGGACGCGAGCGCGCGAGGTGAAGGCTATGG 506  
Qy 542 ACACGCGCGGCTCGCGCGCTGCTCCGCGCTACGTGTCATCCG-TTGGCGACGAGGG 600  
Db 507 ACATTTGTGCGCGTCCGCGCGCTGCTCCGCGCTACGTGTCATCCGCTTGCAGCGAGGG 566  
Qy 601 ATGACTGTGCGCTCGCGCGCGCTGTACGAGGAGAAAGGTGGAGTCTTTGTCAAGCTG 660  
Db 567 ATGACTGTGCGCTCGCGCGCGCTGTACGAGGAGAAAGGTGGAGTCTTTGTCAAGCTG 626  
Qy 661 CGCAGCTCAAGCGCTGTGCGAGGAGAGCTCTGTTCCACTTCTATCTGCGCGTCAATT 720  
Db 627 CGCAGCTCAAGCGCTGTGCGAGGAGAGCTCTGTTCCACTTCTATCTGCGCGTCAATT 686  
Qy 721 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 687 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746  
Qy 781 AAGCTGTGGTACTTTGCCCATGTGCGAGGCTGATGGGTGATTTGACCAACAATTTCTG 840  
Db 747 AAGCTGTGGTACTTTGCCCATGTGCGAGGCTGATGGGTGATTTGACCAACAATTTCTG 806  
Qy 841 GCGGGCTCTGCTGCTGATCAAGTTCTGTTCTGATTTTCCGAGACTCTCTCATAGC 900  
Db 807 GCGGGCTCTGCTGCTGATCAAGTTTGTCTGATTTTCCAGAGACTCTCTCAGAGC 866  
Qy 901 CAGCTTTGGTCTTTTAAAGAGGATCTCTGGCA-TTTCGATTGATAACAAAG 955  
Db 867 CAGCTTTGGTCTTTTAAAGAGGATCTCTGGCATTTTCGATTGATTAACAAAG 922

RESULT 6  
AX513578  
LOCUS  
DEFINITION  
Sequence 15 from Patent WO02059324.  
ACCESSION  
AX513578  
VERSION  
AX513578.1  
KEYWORDS  
Zea mays  
SOURCE  
Zea mays  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1  
Shi, J., Beach, L.R., Wang, H., Rafaleki, J.A. and Cahoon, R.E.  
Novel inositol polyphosphate kinase genes and uses thereof  
Patent: WO 02059324-A 15 01-AUG-2002;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
Location/Qualifiers

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES

source 1. .899  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"  
89\_424  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52958.1"  
/db\_xref="GI:23559680"  
/translation="MPDLHPPEHQVAGHRASAKGLPLIDSGSLFKYKPLQAGDRGEHE  
VAFVEAFSAHAAPVARDITFFPRFHGTRLLPTEAQGPGEHPHYLVLDLLAGFAPCV  
ADIKIGAIT"  
Query Match 51.0%; Score 686; DB 6; Length 899;  
Best Local Similarity 88.7%; Pred. No. 7.5e-98;  
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;  
Qy 2 CACGAGTCAAGTCCGCTCACCCTCGCGCCATAGTCCCTTCCCATACCATGTCGGACC 61  
Db 45 CACCATCGCCACCGCTCACCCTTGTCTCCATAG-----TCCCATACCATGTCGGACC 98  
Qy 62 TCCACCGCGCGAGACCAAGTCCGCGCCACCGCGCTCCGCGAGCAAGCTGGCGCCGC 121  
Db 99 TCCACCGCGCGAGACCAAGTCCGCGCTCACCAGCGCTCCGCGAGCAAGCTGGCGCCGC 158  
Qy 122 TCATCGAGCGCTCGCGCTCTTTCTACAAGCGCTCCAGCGCGGAGCAAGCTGGGGAGCAG 181  
Db 159 TCATCGAGCGCTCGCGCTCTTTCTACAAGCGCTCCAGCGCGGAGCAAGCTGGGGAGCAG 218  
Qy 182 AGGTGCGCTTCTATGAGCGGTTCCTCGGCCACGCGCGCTCCGCGCCGCGCATCCGAGACA 241  
Db 219 AGGTGCGCTTCTATGAGCGGTTCCTCGGCCACGCGCGCTCCGCGCCGCGCATCCGAGACA 278  
Qy 242 CTTCTTCCCGCGGTTCACGCGACGAGCTCTCTCCACCGAGCGCGAGCCCGCGGAGC 301  
Db 279 CTTCTTCCCGCGGTTCACGCGACGAGCTCTCTCCACCGAGCGCGAGCCCGGGAGC 338  
Qy 302 CGCATCTCTCACTCGTCTCGACGACTCTCTCGCGGGTTTTCAGCGCGCTCTGCTGCGAG 361  
Db 339 CGCATCTCTCACTCGTCTCTCGACGACTCTCTCGCGGGTTTTCAGCGCGCTCTGCTGCGAG 398  
Qy 362 ACATCAAGATCGGCGCCATCAAGTGGCCA CCGAGTTTCGCGGAGCCCTACATCGCCAAAGT 421  
Db 399 ACATCAAGATCGTGCATACAGT----- 422  
Qy 422 GCCTCGCCAAAGACCGCGGAGCAAGAGCTTCTGCTCGGATTCGCGCTCTCCGCGTCC 481  
Db 423 -----GACCATGAGCGATCTGCTCGGATTTCCAGCTCTCCGCGTCC 463  
Qy 482 GAGTCGTGCGGCCCGAGGGCGCGGTGTCGCGAGCGAGCGCGCGAGGTGAAGGCGCATGG 541  
Db 464 GAGTCGTGCGGCCCGAGGGCGCGGTGTCGCGAGCGAGCGCGCTCGAGGTGAAGGCTATGG 523  
Qy 542 ACACCGCGGGTTCGCGCGCTGCTTCGCGCGCTACGTGTCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 524 ACATTTGTCGGCTTCGCGCGCTGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583  
Qy 601 ATGAGTCTGTCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTTGTACAGCTG 660  
Db 584 ATGAGTCTGTCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTTGTACAGCTG 643  
Qy 661 CGCAGCTCAAGGCGGTGTTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 720  
Db 644 CGCAGCTCAAGGCGGTGTTTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 703  
Qy 721 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 704 CTTCTGGGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763  
Qy 781 AAGCTGTGGACTTTTCCCATGTCGCGAGGCTGTCGCGAGGCTGATGGGGTGAAGCTCTG 840  
Db 764 AAGCTGTGGACTTTTCCCATGTCGCGAGGCTGATGGGGTGAAGCTGATTTGACCAACAATTTCTG 823

```

Qy 841 GCGGGCTCTGCTGATCAAGTTCGTTTTCACATGTTCCGAGACTCCTCATACG 900
Db 824 GCGGGCTCTGCTAGCTGATCAAGTTGTTTCTGACATTGTTCCAGAGACTCCTTAGACG 883

Qy 901 CAGC 904
    |||
Db 884 CAGC 887

RESULT 7
AP005749
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:OSUNBa0047A17.
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzeae; Oryza.
1
Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSUNBa0047A17
2 (bases 1 to 149142)
Sasaki, T., Matsumoto, T. and Katayose, Y.
Direct Submission
Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Teukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://xgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jun 28, 2004 this sequence version replaced gi:34740257.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM
(http://opal.biology.gatech.edu/glimmer/glmr.fork.html), RiceHMM
(http://opal.tigr.org/tdb/glimmer/glmr.fork.html), SplicePredictor
(http://xgp.dna.affrc.go.jp/RiceHMM/) - SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/html/docs/sim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLAST NonRedundant Protein
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
regions were searched against NCBI NonRedundant Protein database
with BLASTp. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from -21M13 to M13rev of the BAC
clone. This sequence of OSUNBa0047A17 clone has an overlap with
OSUNB0003H22 clone (DBJ: AP005842) at 5' end and an overlap with
P015B12 (DBJ: AP004772) at 3' end. Detailed information on
overlap and assembly quality together with annotation of this entry
is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1. .149142

```

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSUNBa0047A17"
join(2637..2780,3298..3360)
/genes="OSUNBa0047A17.1"
join(2637..2780,3298..3360)
/genes="OSUNBa0047A17.1"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(5864..9107)
/genes="OSUNBa0047A17.2"
complement(join(5864..6755,8843..9107))
/genes="OSUNBa0047A17.2"
/notes="supported by full-length cDNA(s) : AK102815"
complement(join(6112..6755,8843..9017))
/genes="OSUNBa0047A17.2"
/notes="Rieske iron-sulfur protein
contains EST(s) : AU165272(C63256),AU068962(C51210)
contains full-length cDNA(s) : AK102815"
/codon_start=1
/product="putative ubiquinol-cytochrome c reductase
iron-sulfur subunit, mitochondrial precursor"
/protein_id="BA026134.1"
/db_xref="GI:49388912"
/translation="MLRVAGRLTTLAWRPAAAGARGPLAGSLPGDDEFRRBPQR
PFAVDSPFAASRGFSSETLPNQDVLSTELPATVSAVKPNSAKIVDEYVNRHYQ
PGDQSKRAPAYFVLGGRTFYASLLRLLVLSMSASKDVLASLVDLSIEPG
TTVTWKRGKPVFIRTEEDINLANSVDIGSLRDQODAEKVNPEWLVIIGVCTHL
GCPLPNAGDFGWFQFCGSHYDISGRIRKGPAPYNLEVTYSFLEENKLLIG"
9904..10750
/genes="OSUNBa0047A17.3"
join(-9904..10067,10732..>10750)
/genes="OSUNBa0047A17.3"
/notes="start and end point are not identified"
join(9904..10067,10732..10750)
/genes="OSUNBa0047A17.3"
/notes="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD26135.1"
/db_xref="GI:49388913"
/translations="MENATVTADLAPWSQQCRPVGEEEGVAAAKKEGRIQKE
EKGGRRRGSGGIRKTM"
complement(11599..12180)
/genes="OSUNBa0047A17.4"
complement(<11599..>12180)
/genes="OSUNBa0047A17.4"
/notes="supported by full-length cDNA(s) : AK109226"
complement(11599..12180)
/genes="OSUNBa0047A17.4"
/notes="contains full-length cDNA(s) : AK109226
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
join(15870..16043,16479..16484)
/genes="OSUNBa0047A17.5"
join(15870..16043,16479..16484)
/genes="OSUNBa0047A17.5"
/notes="hypothetical ORF
predicted by FGENESH
this category is not included in IRGSP standard"
complement(18932..25995)
/genes="OSUNBa0047A17.6"
complement(join(18932..19648,19740..19836,20081..20194,
20280..20444,20690..20827,20923..20999,21093..21207,
21277..21382,21703..21764,21859..21993,22149..22331,
22880..22936,23369..23417,23504..23652,24724..24771,
25582..25995))
/genes="OSUNBa0047A17.6"
/notes="supported by full-length cDNA(s) : AK120851"

```

FEATURES  
source



Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC clone: P0415B12  
Published Only in Database (2002)  
2 (bases 1 to 157419)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (20-FEB-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
On Jun 28, 2004 this sequence version replaced gi:37497094.  
Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerHMM (http://www.tigr.org/tdb/glimmer/glmr form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologs of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no.  
A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative', and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.  
The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0415B12 clone has an overlap with OSJNB0047A17 clone (DBJ: AP005749) at 5' end and an overlap with P0458805 clone (DBJ: AP004777) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.  
Location/Qualifiers  
1..157419  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="P0415B12"  
complement(4313..5807)  
/gene="P0415B12.1"  
complement(join(4313..4547,5568..>5807))  
/gene="P0415B12.1"  
/note="supported by full-length cDNA(s): AK121467"  
complement(4313..5807)  
/gene="P0415B12.1"  
/note="contains full-length cDNA(s): AK121467 non-coding transcript probably inactive due to including stop codon(s) in CDS probably inactive due to too long 3'UTR in CDS"  
6065..7404  
/gene="P0415B12.2"  
join(<6065..6067,6850..>7404)

/genes="P0415B12.2"  
/notes="start and end point are not identified"  
join(6065..6067,6850..7404)  
/gene="P0415B12.2"  
/note="predicted by FGENESH etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD25374.1"  
/db\_xref="GI:49388256"  
/translation="MNSRQIPLVYPRHEIVARLAKEDGGHQEQOQTQEVKGDGHD SKNTLPSPASPAATAPASALATIVASELPEEQNHALLPTVKAQILSAPKILSS SDTSTVTITGTAAENSEGTAMVAATVCMKSGLEKPTNPRVTRHVHVGSLGRLC TVQEGVGEGVLEGRGSCQITVPC"  
complement(8669..15038)  
/genes="P0415B12.3"  
complement(join(8669..9060,9161..9334,9438..9569,9640..9738,9929..10000,10080..10136,10206..10379,10460..10554,10727..10826,10921..11073,11145..112275,12870..12980,13078..13240,13613..13707,13793..14053,14178..14331,14775..15038))  
/genes="P0415B12.3"  
/note="supported by full-length cDNA(s): AK121782"  
complement(join(8929..9060,9161..9334,9438..9569,9640..9738,9929..10000,10080..10136,10206..10379,10460..10554,10727..10826,10921..11073,11145..112275,12870..12980,13078..13240,13613..13707,13793..14053,14178..14331))  
/genes="P0415B12.3"  
/notes="contains EST(s): AU082722(R4038), AU032417(R4038) contains full-length cDNA(s): AK121782"  
/codon\_start=1  
/product="gamma-tubulin complex component 5-like"  
/protein\_id="BAD25375.1"  
/db\_xref="GI:49388257"  
/translation="MVLPLNAGLMEADSGSFIKQLQSVSDGLPHAAPVPELSTQE HELVSVFEVQFGDTVLLYWDKTPGCEKAGIYVSHLSQTSRLAVKLPFAATCL KQVLFVGRVRSQCHGPTLTSAPASSVDWMLRLKALKEEQFLTSVETITLLGL TDSMSIACSAEHLVQVQGVAPDAFWNSGAMASSEVAHVAVNHFKKLEVECLD GEGYPHMLAVIAGTLLPYQLCLDSWLYDGLDDPNEEMFFYANKVATIDQPAFWM SYMRVGRPTDSSSLADNESIRKELINQETATAAALLKSSNOGCADILCPVFLKD IARALISAGKSFQVHQVQTHRIQREVIFHEFNIDQHGNTISQKFRPDTSSIRIQD KREDITBESTQFGNACKMDFLTLSFSLICSLGLENGHDVDDYLRKLDNAPVFN PGHDVSKSVQTEVEVGENSGENKTLKLRDATSGRDYGMETKLNKAVMRDPTFV KXIFGDRDSSGTIPGDTLSTVFPRLDGTGKGFQDDSEVIROEDDRRTEALY TPFTLPCVENVPLSELPLQKDSLASALAKFQMSLRDPLQVPIQECLSKCI KRQVDHIGKQILSKMGDWRMLDFVIRAIYLLIGSGDMLQQFLVTITFDKLDKGNPD DPELNTLLQESIRNSADKMLTAPDSLVSVLAKHTRNDEETTSIRKGRQAQFGIE ALDVNFTYKVSWPLDLIVNTEALKKYNQVMAFLKVKRAKFIIDETRKWKKGGST MNFKOHLIVEOKLHFVDAPFOVMDRVVHSAWTELCGWSATTILDEVMEVHAYL SSTORQCFVADKLWALIASRVKLTILGLALDFHNIETGLTGTGTPAPKARCEMEVD R IERQDFCVFLRLILSFKLVNGVGHFPHADLVTRINTYHMYMSDGSFSAIPGSRPR"  
complement(16514..23105)  
/gene="P0415B12.4"  
complement(join(16514..17069,17267..17644,18400..18594,18674..18751,18816..19001,19077..19307,19386..19739,19837..20034,20118..20288,20405..20512,20611..20829,20914..21120,21227..21445,21564..21636,21998..22166,22901..23105))  
/gene="P0415B12.4"  
/note="supported by full-length cDNA(s): AK065061"  
complement(join(16881..17069,17267..17644,18400..18594,18674..18751,18816..19001,19077..19307,19386..19739,19837..20034,20118..20288,20405..20512,20611..20829,20914..21120,21227..21445,21564..21636,21998..22150))  
/gene="P0415B12.4"  
/note="contains EST(s): AU055859(S20105), AU085756(C52347), C27605(C52347), AU078722(S20105) contains full-length cDNA(s): AK065061, AK063159"  
/codon\_start=1  
/product="RNA binding protein Rp120"  
/protein\_id="BAD25376.1"





Yoshino, M. and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 1570)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Ikeda, Y., Ikeda, Y., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurohata, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

TITLE

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Ooka, H., Kojima, K., Shishiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurohata, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ooka, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers

1..1570  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cui="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J023018G11"

ORIGIN

Query Match 40.5%; Score 544.8; DB 8; Length 1570;  
Best Local Similarity 77.4%; Pred. No. 9e-76;

| Matches | 717; | Conservative   | 0; | Mismatches | 192; | Indels | 17; | Gaps | 4; |
|---------|------|--|----|------------|------|--------|-----|------|----|
| Qy      | 13   | TCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGACACCTCCACCCCGCG   |    | 72         |      |        |     |      |    |
| Db      | 344  | TCACACACCCCGCGCGCCCATAGTCCCTTCCCATACCATGTCGACACCTCCACCCCGCG  |    | 401        |      |        |     |      |    |
| Qy      | 73   | GAGCACCACAGTTCGCGCGCCACCGCGCTCCGCCAGCAAGCTGGGCGCGCTCATCGAGCG |    | 132        |      |        |     |      |    |
| Db      | 402  | GAGCACCACAGTTCGCGCGCCACCGCGCTCCGCCAGCAAGCTGGGCGCGCTCATCGAGCG |    | 461        |      |        |     |      |    |
| Qy      | 133  | TCGCGCTCTTTCACAAAGCGCTCCAGCGCGCGACCGTGGGAGACAGAGTTCGCTTC     |    | 192        |      |        |     |      |    |
| Db      | 462  | GAGCGGCTCTTTCACAAAGCGCTCCAGCGCGCGACCGTGGGAGACAGAGTTCGCTTC    |    | 521        |      |        |     |      |    |
| Qy      | 193  | TATGAGCGCTTCGCGCGCCACCGCGCTCCGCCAGCAAGCTGGGCGCGCTCATCGAGCG   |    | 252        |      |        |     |      |    |
| Db      | 522  | TATGAGCGCTTCGCGCGCCACCGCGCTCCGCCAGCAAGCTGGGCGCGCTCATCGAGCG   |    | 581        |      |        |     |      |    |
| Qy      | 253  | CGGTTCCACGCGCAGCGCTTCCTCCCGCCAGCGCGCGCTCCCGCGCGCTTCCTCCCG    |    | 309        |      |        |     |      |    |
| Db      | 582  | CGGTTCCACGCGCAGCGCTTCCTCCCGCCAGCGCGCGCTCCCGCGCGCTTCCTCCCG    |    | 641        |      |        |     |      |    |
| Qy      | 310  | CACCTCGCTTCGAGACCTTCCTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG   |    | 369        |      |        |     |      |    |
| Db      | 642  | CACATCGCTTCGAGACCTTCCTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG   |    | 701        |      |        |     |      |    |
| Qy      | 370  | ATCGGCGCTTCGAGACCTTCCTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG   |    | 429        |      |        |     |      |    |
| Db      | 702  | ATCGGCGCTTCGAGACCTTCCTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG   |    | 761        |      |        |     |      |    |
| Qy      | 430  | AAGGACCGCGCGCTTCCTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG       |    | 489        |      |        |     |      |    |
| Db      | 762  | AAGGACCGCGCGCTTCCTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG       |    | 821        |      |        |     |      |    |
| Qy      | 490  | ---GGCGCGCGCGCTTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG         |    | 546        |      |        |     |      |    |
| Db      | 822  | GATGCGCGCGCGCTTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG          |    | 881        |      |        |     |      |    |
| Qy      | 547  | GCAGCGCTTCGCGCGCTTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG       |    | 606        |      |        |     |      |    |
| Db      | 882  | GCAGCGCTTCGCGCGCTTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG       |    | 941        |      |        |     |      |    |
| Qy      | 607  | TGTCGCTTCGCGCGCTTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG        |    | 666        |      |        |     |      |    |
| Db      | 942  | TGTCGCTTCGCGCGCTTCGCGGGGTTCAGCGCGCTTCGCTCGGAGACATCAAG        |    | 1001       |      |        |     |      |    |
| Qy      | 667  | CTCAAGCGCTTCGAGGAGACAGCTCTGTTCCATCTTCTACTCGCGCTCGATTTCTG     |    | 726        |      |        |     |      |    |
| Db      | 1002 | CTCAAGCGCTTCGAGGAGACAGCTCTGTTCCATCTTCTACTCGCGCTCGATTTCTG     |    | 1061       |      |        |     |      |    |
| Qy      | 727  | GGCT-----ATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG       |    | 777        |      |        |     |      |    |
| Db      | 1062 | GGCT-----ATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG       |    | 1121       |      |        |     |      |    |
| Qy      | 778  | GTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG       |    | 837        |      |        |     |      |    |
| Db      | 1122 | GTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG       |    | 1181       |      |        |     |      |    |
| Qy      | 838  | CTGGCGCGCTTCGCTCGCTGATCAAGTTCGTTCTTCTGACATTTCTCGGAGACTCTCAT  |    | 897        |      |        |     |      |    |
| Db      | 1182 | CTGGCGCGCTTCGCTCGCTGATCAAGTTCGTTCTTCTGACATTTCTCGGAGACTCTCAT  |    | 1241       |      |        |     |      |    |
| Qy      | 898  | ACGCGAGCTTTGGGCTCTTCTTAAAG 923                               |    |            |      |        |     |      |    |
| Db      | 1242 | AAGCGCTTTTCTGATCATTTCTTGAAG 1267                             |    |            |      |        |     |      |    |

RESULT 10

AX513580

LOCUS

DEFINITION

AX513580

ACCESSION

VERSION

AX513580

Sequence

17 from Patent

WO02059324.

AX513580

GI:23559681

linear

DNA

643 bp

PAT 05-OCT-2002





```
Db 1 CTAAAGCATGGTTGGAGAGCAGACTCTGTTCACACTTCTACTCGGGTTCGATTTCTTG 60
Qy 727 GGCTATGATGCTGCTGAGTGCAGAGCGAGAGTGGGGTGGGGTGGAGGTCGAAGCTG 786
Db 61 GGCTATGATGCTGCTGAGTGCAGAGCGAGAGTGGGGTGGGGTGAACAGTGAAGCTG 120
Qy 787 GTGACTTTGGCCATGTGGCGAGGGTGAATGGGGT--GATTGACCAACATTTCTGGGCG 844
Db 121 GTGACTTTGGCCATGTGGCGAGGGTGAATGGGGTGAATTTGACCAACATTTCTGGGCG 180
Qy 845 GGCTGCTGCTGATCAAGTTC--GTTTCTGACATTTGTCGGAGACTCTCTATAG-C 901
Db 181 AGCTGCTGATGATCAAGTTCGGTTCTTTGACATTTGTCAGANACTCTTAGAGGCC 240
Qy 902 AGCCTTTGGTCTCTTAAGAGA 925
Db 241 AGCCTTTGGTCTCTCTTAAAAA 264

RESULT 13
AX513574
LOCUS AX513574 1195 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 11 from Patent WO02059324.
ACCESSION AX513574
VERSION AX513574.1 GI:23559674
KEYWORDS Eucalyptus grandis
SOURCE Eucalyptus grandis
ORGANISM Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Myrtales; Myrtaceae; Eucalyptus.
1
REFERENCE Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.B.
AUTHORS Novel inositol polyphosphate kinase genes and uses thereof
TITLE Patent: WO 02059324-A 11 01-AUG-2002;
JOURNAL PIONEER HT-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1..1195
/organism="Eucalyptus grandis"
/mol_type="unassigned DNA"
/db_xref="taxon:71139"
116..1048
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52956.1"
/db_xref="GI:23559675"
/translation="MLKVPDQVAGHRGDKGLPLVDSDGRFYKPLQSDHGRGTEVA
FYGSFYNTPEIPGHIRKFPFAPHTGTIEASDGSPHPLVLEDLVSRGTPSLMDIK
TGSRTWYPEASEEYIOKLEDRNSTSVSLGFRISGLRVYONSEAGFWQPEKKVYVSF
NADGVRSLRPFVSNLSLGNVDPLCLYASKYVCHRGGILAQLLQLEKFEFVQTNH
FYSCSLIILYDRESALDGAHPKVKLVDFAHVMDGHVIDHNFVLGLGCLSVIKFIRDIA
DEDNKAACEYNLGLKENGIFYKSSTPELDHEAC"

CDS
ORIGIN
Query Match 13.7%; Score 184.6; DB 6; Length 1195;
Best Local Similarity 55.6%; Pred. No. 2.3e-19;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;
Qy 56 CCGACTCCACCGCGGAGCAGCAGCAAGTCGCGGCGACCGCGCTCCCGCAGCAAGCTGG 115
Db 114 CCATGCTCAAGTCCGGATCATCAAGTCGCGGTACCGGGAGACGGGGAAAGCTGG 173
Qy 116 GCCCGCTCATGAAGCTCGCGCTCTTCTTCAAGCGCTCAGCGCGCGGACCGCTGGGG 175
Db 174 GGCACACTGGTGAATTCGGGCGCTCTATAGCTCTCAGAGCGATCATCGCGAG 233
Qy 176 AGCAGAGTGCCTCTTATAGGGCTTCTCGCCACGCGCGCTCCCGCCCGCATCC 235
Db 234 ACACGGAAGTGGCCCTTTTACGAGTCAATCTATTTCAATACCGAGATCCCAATTC 293
Qy 236 GAGACACCTCTTTCGCCGGTTCCACGCGAGCGGACTCTCCACCGAGGCGAGCCCG 295
Db 294 G---CAATTTCTTCTCGCTTTCAGGAACTAAGACTAT---TGAGGCGTCTGATGAT 347
```

```
Qy 296 GGGAGCGGCATCCTCACTCGTCTCGACACCTCTCTCGCGGGTTTTCAGGCGCCCTCGC 355
Db 348 CGGGTCTCTCACTCCTACCTCGTCTCGAGGATCTCGTCTCGGGTCGACGAACCACTC 407
Qy 356 TCGAGACATCAAGATCGGCGCATCACGTGGCCACCGAGTTCCGCGGAGCCCTACATCG 415
Db 408 TCATGGACATCAAGATCGGATCCAGAACATGTGTATCGGAGGCTCTGAGGAGTACATCC 467
Qy 416 CCAAGTGCCTCGCAGGACCGGCGGACACGAGCGTTCTGCTCGGATTCGGGCTCCG 475
Db 468 AAAAGTGTCTTAGAGAAAGATCGAAATAGCAACAGCGTTTCATTGGGTTTATGGAATTC 527
Qy 476 GCGTCCGAGTCTCGG---CCCGAGGGCGCGGTGTGGCGGACGCGGCGCGGAGGTGA 532
Db 528 GGCTAAGGTATATCAAAATAGCGAGCTGGATTTTGGCAACCTGAGAGAGGTTGTT 587
Qy 533 AGGCCATGGAACACCGCGCGCTCGCGCGTGTGTCTCGGCGCTAGCTGTCTATCC----- 585
Db 588 ATAGCTTTAATGCGGACGCGTGTGAGTGGCTCTGAGGAAAGTTGTTTCTTCCAACTGT 647
Qy 586 -----GTTGCGGAGGAGGATGGACTGTGCGTCTCGCGCGGCTGTACGAGGAAAG 640
Db 648 CTCGGGTCCAAATGTGGATCCGGATTTGTGTATGCATCAAAAGTTTACTGTCCCGGG 707
Qy 641 GTGAGTCTTGTCAACAGCTCGCGGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTGTTCC 700
Db 708 GTGGAATTTGGCAATTCCTTCAGCTGAAGAAATGGTTGAGTTTCAGACGAATATC 767
Qy 701 ACTTCTACTCGGCGTCAATTTCTTCTGGGCTATGATGCTGTGAGTGCAGAGCGCGAG 760
Db 768 ACTTCTATTCTGTTCACTCAATTTCTATATGACAGGAGTCTGCTTGGACGCG--- 823
Qy 761 ATGGGGTGGGTGACGCTGAAGCTGTGACTTTGCCATTTGCCATGTGCGGAGGATCGGG 820
Db 824 --TGTGCACACCCGAAAGTTAAACTGGTGGACTTTGCATGTGATGATGATGCGCACGCG 881
Qy 821 TGATTGACCAACTTCTCGGCGGCTCTGCTGCTGATCAAGTTCTGTTTCTGACATTG 880
Db 882 TGATCGATCAAACTTCTTGGTGGCTCTGTTCTGTAATCAAGTTTATAGTGACATTG 941
Qy 881 TTCGGA 887
Db 942 CTGATGA 948

RESULT 14
LUS310150
LOCUS LUS310150 25054 bp DNA linear PLN 24-JAN-2002
DEFINITION Linum usitatissimum variety Bombay Ngc-D, Ngc-A and Ngc-B genes.
ACCESSION AJ310150
VERSION AJ310150.1 GI:13509206
KEYWORDS Ngc-A gene; Ngc-A protein; Ngc-B gene; Ngc-B protein; Ngc-D gene; Ngc-D protein.
SOURCE Linum usitatissimum (flax)
ORGANISM Linum usitatissimum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Linaceae; Linum.
1
REFERENCE Dodds, P.N., Lawrence, G.J. and Ellis, J.G.
AUTHORS Contrasting modes of evolution acting on the complex N locus for
TITLE rust resistance in flax
JOURNAL Plant J. 27 (5), 439-453 (2001)
MEDLINE 21461288
PUBMED 11576428
REFERENCE 2 (bases 1 to 25054)
AUTHORS Dodds, P.N.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2001) Dodds P.N., Plant Industry, CSIRO, GPO Box
1600, Canberra, ACT 2601, Australia
FEATURES
Location/Qualifiers
1..25054
source
```

```

/organism="Linum usitatissimum"
/mol_type="genomic. DNA"
/variety="Bombay"
/db_xref="taxon:4006"
/haplotype="N"
join(2881. .3431,3533. .4652,4770. .5045,5140. .6519)
/gene="Ngc-D"
join(2881. .3431,3533. .4652,4770. .5045,5140. .6519)
/gene="Ngc-D"
/note="N rust resistance gene: TIR-NBS-LRR protein"
/codon_start=1
/product="Ngc-D protein"
/protein_id="CAC35321.1"
/db_xref="GI:13509207"
/db_xref="GOA:Q9AR4"
/db_xref="UniProt/TREMBL:Q9AR4"
/translation="MRSDDSGSDSFHSCSSADLTLPPLPSGEVEVFLSRGPDVRK
TFADHLVTSVRSKIRTRDEEGKQGETIGPSLIQAITESKIYIPILTONYASSKWC
LOELAKVNCWKGCGEAKGQHIIIPVYFMDPRDVRHPDPSGPKESPEOHNLKHPDET
VLEWKEALQVCKGKGYHVTFSDGHGSIIDKILTEVHLGAGNAYALVDELVDGDSRV
DEWGLLNDSDTSEKIIGHGMGLGKTTAKAVYDKVSTKFCYCFLENIROTLSB
KNGVILQNKIISGILRKDFNEAKNADGIRIIRDRVCRHKLILVLDVDKRFQDDV
LQKNFTNSRFLITTRDARGLELLQYKMFELQMSPDHSLTILFNKHAFDVDCPPK
DYALSKSFQAAAGLPYIKVIGSLFRMDKIFWEKLEBFKLSPTKVQVERKISY
NELTHNEQIFLDIACYFIGSQKIYPIFMWDCDFPESTIRSLQIRLSRIK
GDVLTFFWHDHIIIDGRVIREKNQNPYKRSIRWSNKDAVNMKHKKGDTCEVSLY
WMEGEDILITNKPEKLTMLRYLKVSNAKLAGDPKVLNLRMLLLESCDPSGVLV
LKLVLRLDHCSDGSDHGMNKLKAVKLSKRCFHLKVPDFSDCGDLFLNF
DCCNRNREVDIGNFKSRFLYISKTITKIKEIGRGLNKLISVGDSSESKEVPAGI
SKLSLEFLALADTSYKDFTEMLPTSLTLLYSNDQKFCPTDSTSNLQRLPNLS
LNLVSLVLDIGEILGELKMLYLSIGRASRIVHLGLNLENLVLQHLRVEGCR
ILRKLPSLIATRLQLLWQDCLVTEINGMGLWESLSHLKVGCSALIGLESLSHM
VKLERLLVGCVLTEPMPSLSMFTKLTSLSCAMPKQFPDLNKLNRLLKVSFCQ
ELIEVPGDALSLEKWLNSGCRSIRKVPDLSGLKXKLTLDVESCIOQLKEVRLGLE
SUELKMGCSIEELPNLSGLNKLRELLKGCIOQLKEVNGLEGLLELTVFEARKRIKA
KYVMSAARYGQLTTSRN"
4553
/gene="Ngc-D"
/standard_names="n-X50"
/note="site of 11 bp insertion in n-X50 rust susceptible
mutant allele"
/evidence=experimental
5608
/gene="Ngc-D"
/standard_names="n-X128A"
/note="site of LUTE transposon insertion in n-X128A rust
susceptible mutant allele"
/evidence=experimental
6397. .6629
/standard_names="n-X139"
/note="233 bp deletion in n-X139 rust susceptible mutant
allele"
/evidence=experimental
9613. .10515
/function="unknown"
/note="ORF
related to human inositol hexakisphosphate kinase 2
protein"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC35322.1"
/db_xref="GI:13509208"
/db_xref="GOA:Q9AR3"
/db_xref="UniProt/TREMBL:Q9AR3"
/translation="MFKVPQHVAGHODINGLLGPLVDDSGRFYKPLQDGRGATEMA
FYTSFAKTIPOHVRFPFPAHQTSIEASDGLHPLHILEDTSLRLHCVMNDIK
IGSRWYPPASOAYTEKCLKDVSSSPFLGFRISGLQVNGKSSSVLKPBRKLL
QNLTADEVLVLKRVSSNPKSDQDCSFAAVYGGSGILAQLELKNWEDOTIYH
FNSCVLMLEYEKTKMVGGEESLGCRAVRUIDFAHVTGNGVIDHNFPLGGLCSLIK
FISELITGDPENSPKSCLODGAH"
join(12668. .13218,13450. .14542,14757. .15032,15186. .16493)
/gene="Ngc-A"
join(12668. .13218,13450. .14542,14757. .15032,15186. .16493)

```

```

/gene="Ngc-A"
/note="N rust resistance gene homolog: TIR-NBS-LRR
protein"
/codon_start=1
/product="Ngc-A protein"
/protein_id="CAC35323.1"
/db_xref="GI:13509209"
/db_xref="GOA:Q9ARC2"
/db_xref="UniProt/TREMBL:Q9ARC2"
/translation="MRSDDSGSDSFHSCSSADLTLPPLPSGEVEVFLSRGPDVRQ
TFADHLVTSVRSKIRTRDEEGKQGETIGPSLIQAITESKIYIPILTONYASSKWC
LOELAKVNCWKGCGEAKGQHIIIPVYFMDPRDVRHPDPSGPKESPEOHNLKHPDET
ILEWKGALQEVGKMGWHISLITQCGAVVDKIITEVHLRANTLATDELVDGDFSV
DEWGLLNDSDTSEKIIGHGMGLGKTTAKAVYDKVSTKFCYCFLENIROTLSB
KNGVILQNKIISGILRKDFNEAKNADGIRIIRDRVCRHKLILVLDVDKRFQDDV
LQKNFTNSRFLITTRDARGLELLQYKMFELQMSPDHSLTILFNKHAFDVDCPPK
DYALSKSFQAAAGLPYIKVIGSLFRMDKIFWEKLEBFKLSPTKVQVERKISY
NELTHNEQIFLDIACYFIGSQKIYPIFMWDCDFPESTIRSLQIRLSRIK
GDVLTFFWHDHIIIDGRVIREKNQNPYKRSIRWSNKDAVNMKHKKGDTCEVSLY
WMEGEDILITNKPEKLTMLRYLKVSNAKLAGDPKVLNLRMLLLESCDPSGVLV
LKLVLRLDHCSDGSDHGMNKLKAVKLSKRCFHLKVPDFSDCGDLFLNF
DCCNRNREVDIGNFKSRFLYISKTITKIKEIGRGLNKLISVGDSSESKEVPAGI
SKLSLEFLALADTSYKDFTEMLPTSLTLLYSNDQKFCPTDSTSNLQRLPNLS
LNLVSLVLDIGEILGELKMLYLSIGRASRIVHLGLNLENLVLQHLRVEGCR
ILRKLPSLIATRLQLLWQDCLVTEINGMGLWESLSHLKVGCSALIGLESLSHM
VKLERLLVGCVLTEPMPSLSMFTKLTSLSCAMPKQFPDLNKLNRLLKVSFCQ
ELIEVPGDALSLEKWLNSGCRSIRKVPDLSGLKXKLTLDVESCIOQLKEVRLGLE
SUELKMGCSIEELPNLSGLNKLRELLKGCIOQLKEVNGLEGLLELTVFEARKRIKA
KYVMSAARYGQLTTSRN"
17600. .18457
/note="unknown"
/note="ORF
related to human inositol hexakisphosphate kinase 2
protein"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC35324.1"
/db_xref="GI:13509210"
/db_xref="GOA:Q9ARC1"
/db_xref="UniProt/TREMBL:Q9ARC1"
/translation="MLKVPQHVAGHODINGLLGPLVDDSGRFYKPLQDGRGATEMA
FYTSFAKTIPOHVRFPFPAHQTSIEASDGLHPLHILEDTSLRLHCVMNDIK
IGSRWYPPASOAYTEKCLKDVSSSPFLGFRISGLQVNGKSSSVLKPBRKLL
QNLTADEVLVLKRVSSNPKSDQDCSFAAVYGGSGILAQLELKNWEDOTIYH
FNSCVLMLEYEKTKMVGGEESLGCRAVRUIDFAHVTGNGVIDHNFPLGGLCSLIK
FISENQLML"
join(20265. .20815,21104. .22202,22415. .22690,22818. .24218)
/gene="Ngc-B"
join(20265. .20815,21104. .22202,22415. .22690,22818. .24218)
/gene="Ngc-B"
/note="N rust resistance gene homolog: TIR-NBS-LRR
protein"
/codon_start=1
/product="Ngc-B protein"
/protein_id="CAC35325.1"
/db_xref="GI:13509211"
/db_xref="GOA:Q9AR40"
/db_xref="UniProt/TREMBL:Q9AR40"
/translation="MRSDDTSSIGSLRSRSDPTLPPLPSGEVEVFLSRGPDVRQ
TFADHLVTSVRSKIRTRDEEGKQGETIGPSLIQAITESKIYIPILTONYASSKWC
LOELAKVNCWKGCGGQHIIIPVYFMDPRDVRHPDPSGPKESPEOHNLKHPDET
ILEWKEALQVCKGKWHINELTQCGAVVDKIITTEPHLRANTLATDELVDGDSV
BEVMEMLNDHSTSERIIGYMGGLGKTTAKAVFNKVSQMFQFCFCDLNIREFLLR
NDGVVNLQNKVSIIDLRKDSQAKNADGVRIIRVRVRHKI FVVLDDIDESFHYDEI
FGKLDGSDTSRFLITTRDARTLELNECKMFGLEEMSHDSLQLSFKHAFQVDPPE
DYALSKSEFIQVAGSLPALKVGSLFPSDKRFWEDKLIEKLALPSAKVQRKLKVS
NETHYNEQIFLDIACLVFAGKVPVPMWDCDLPASTLRTLVQRLVRMDNDKRF
WHDHIDRGLGRTI VREENSQNPYKRSIRWSNKDAIDILNREGNDCVCLRVDMKGE
YALNTRKFKQSRFLRFLVNLGDSGNFPLNLRVLRVYRGDPSGLNLCVMKGVIL
ELDCGYVTHWKGHNEIKAKCLKAVNLTCGILEKVPDLSTCRGLELCKHCKCOWMR
GELDITGFKDLKVLINDQTEITTLKGEVESIQNLQQLDVGSRGLEVPAGLSLSLE
FLDITSVKHDSEVEMLPNGLKULVSSFSLSALPSSLIKLDICDSNRLQRLPNLASYTN
LTRLHLEVGTHETIPGLGKLKLESLSICNAPNLDNLDGLNLLKELALERCPCILG
KLPSLAELTKLHKVIRWCDVLGEIYGLNGLGDSLHLDISWCPLRTVMDLHLSLLKL

```

CDS

gene

CDS

GTLLVSGFELTNILPLSLIYVTKLRTLEVRSSQPLDNLNKLNRDLTITGCRELIEIA  
GLHTLESLELSMERCPSVKRLDIAGLIKLTIHICLTOLTEIRGGGLESLOMLFM  
SGCOSIHELNLGSLXNLKYPSELKECQLKEVNGLEBLEWLDFTDRRLKLYLLKLV  
SRKGQLVTSARRALFINN"

ORIGIN

Query Match 13.5%; Score 181.6; DB 8; Length 25054;  
Best Local Similarity 55.4%; Pred. No. 4.9e-19;  
Matches 474; Conservative 0; Mismatches 349; Indels 33; Gaps 5;  
Qy 50 CCATGTCGACCTCCACCCGCGGAGCACCAAGTGGCGGCCACCGCGCTCCGCCAGCA 109  
Db |||||  
Qy 9605 CCAAAGCCATGTTCAAAGTTCGGGAACATCAGGTGCTGGACATCAAGCCATCAACGGCC 9664  
Db |||||  
Qy 110 AGCTGGGCGCGCTCATCGACGGCTCCGGCTCTTCTACAAGCGCTCCAGCGCGCGGACC 169  
Db |||||  
Qy 9665 TCCTCGGTCTCTTTGTAGACGATTCAGGGCGATTTCTACAAGCCACTTCAGGGCGATGAAC 9724  
Db |||||  
Qy 170 GTGGGGACACGAGTCCGCTTCTATGAGCGTTCTCGGCCACCGCGCTCCGGGCC 229  
Db |||||  
Qy 9725 GTGGGGACACGAGATGGCATTTCTACATCTTTTCTGCTCACAAGACGATTCACAC 9784  
Db |||||  
Qy 230 GCATCCGAGACACCTTTCTCCCGGTTTCCAGGCACGCGACTCTCTCCCAACCGAGGCGC 289  
Db |||||  
Qy 9785 AGTCCG---GAGGTTTTTCCCGGCTTCCATGGCACTCAAGTATAGAGGCTTCTGATG 9841  
Db |||||  
Qy 290 AGCCCGGGAGCGGATCCTCACTGCTGCTCGACGACTCTCTCGCGGGTTTCAGGGCG 349  
Db |||||  
Qy 9842 GATCTGG---CTCCACCGCACCTCATCTTTAGAGATCTCACGTGCACTCGGTTGCATC 9898  
Db |||||  
Qy 350 CCTGGTCGACACATCAAGATCGCGCCATCACTGTCGCCACCGAGTTTCGCGGAGCGCT 409  
Db |||||  
Qy 9899 CATGTCATGACATCAAGATTGGTTCAGACATGTTTACCGAGGCTTCTCAAGCCT 9958  
Db |||||  
Qy 410 ACATCGCCAAAGTCCCTCCCAAGCACCGCGGACCAAGAGGTTCTGCTCGGATTCGCGC 469  
Db |||||  
Qy 9959 ACATCGAATAATGCTTGAAGAAGGATGTGAATCAAGCAGTCCCTTTCTCGGGTTTAGGA 10018  
Db |||||  
Qy 470 TCTCGGCGTCCGAGTCTGTCGCGCCGA-----GGGCGCGTGTGGCGGACGG 517  
Db |||||  
Qy 10019 TATCCGATTCGAGGTTTACGGTTAACGCAAGAATCATCTGAGGTGCTTTGAAGCCTG 10078  
Db |||||  
Qy 518 AGCGCCCGAGGTGAAGCCATGACACCGCGCGTCCGCGGTGCTCGCGGTACG 577  
Db |||||  
Qy 10079 AGAGAAAGCTTCCAGAACTTACTGCTGATGAAGTTAGGCTGGTGTAAACGGTTTG 10138  
Db |||||  
Qy 578 TGTCAATCGTTGCCGA---CGAGGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGAG 634  
Db |||||  
Qy 10139 TTTCTCTCAACCCGAAGTCGACCAACCGGATTTGTTCTGTTGACGCGTCTTTTACGGTG 10198  
Db |||||  
Qy 635 GAAAGGTGGAGTCTTGTCAACGCTGCGGAGCTCAAGCGTGGTTCAGAGAGGAGTCTC 694  
Db |||||  
Qy 10199 GTTCAATGGGATCTTGGCTCAGTTGCTGAGCTGAAGCGGTGGTTTGAAGATCAGACCA 10258  
Db |||||  
Qy 695 TGTTCACCTTCTACTCGGCGTCAATCTTCTGGGCTATGATGCTCT-----G 742  
Db |||||  
Qy 10259 TCTATCATTTCAACTTTGTTCTGTCTTATGCTGTATGAGAGGAGAACGAGATGG 10318  
Db |||||  
Qy 743 CAGTCGACGAGCGGAGATGGGGTGGGGTGACCGTGAAAGCTGTGAGACTTTTGCCTATG 802  
Db |||||  
Qy 10319 TGAATGGAGGAGAGAGATTTGGGTTGTGCTGCGGTGAAGCTCATCGATTTGCTCATG 10378  
Db |||||  
Qy 803 TGGCGAGGGTGAATGGGTGATGACCAACACTTCTCGGGCGGGCTCTGCTGCTGATCA 862  
Db |||||  
Qy 10379 TGACGGAGGGAATGGCGTTATCGATCACAACCTTCTTGGAGGGCTCTGCTCAITGATAA 10438  
Db |||||  
Qy 863 AGTTCGTTCTGACAT 878  
Db |||||  
Qy 10439 AGTTCATTTCCGAGAT 10454  
Db |||||

RESULT 15  
AX513572

LOCUS AX513572 1105 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 9 from Patent WO02059324.  
ACCESSION AX513572  
VERSION AX513572.1 GI:23559672  
KEYWORDS Glycine max (soybean)  
SOURCE Glycine max  
ORGANISM Glycine max  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1  
REFERENCE  
AUTHORS Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.  
TITLE Novel inositol polyphosphate kinase genes and uses thereof  
JOURNAL Patent: WO 02059324-A 9 01-AUG-2002, (US)  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
Location/Qualifiers  
1..1105  
/organism="Glycine max"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3847"  
12..851  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD52955.1"  
/db\_xref="GI:23559673"  
translation="MLKIPHQVAGHAKDGIILGPLVDDRFKFKYKPLQTNKDDTTRGS  
TELSFTSLAAAHDSYIRSFFAFHGTLLDASDGGPHPLVLELLCGYSKPSVM  
DHVGRTRHLDGSEDIYCKLKKDRSSSLPLGFRISGVKDSISSMEPTKSLQCLUS  
AHGVALVANKFVSSNNIHDHDPDCAFATVYGAFLERLQKLDWFEVQTVVHYFSC  
SLVAVYEKDLGKGKATNPLVKLVDFAHVVDNGVIDHNFLGGLCSFKPLKDLILAVAC  
LHK"  
ORIGIN  
Query Match 13.3%; Score 179; DB 6; Length 1105;  
Best Local Similarity 55.6%; Pred. No. 1.7e-18;  
Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;  
Qy 61 CTCACCCGCGGAGCACCAAGTCGCGGCACCGCGCTCCGACCAAGCTGGGCGCCG 120  
Db |||||  
Qy 15 CTCAAGATCCCGAGCACCAAGTGGCGGCGCACAAAGCCACGGAATCTCTGGGCCCCA 74  
Db |||||  
Qy 121 CTATCGACCGCTCCGCGCTTCTTACAAGCGCTCCAGCGCGCGCGACCGTGGGGAGCAC 180  
Db |||||  
Qy 75 CTGTCGACGATTTGAAAAATTTCTACAGCCCTCCAGACCAACAAAGACGACACACC 134  
Db |||||  
Qy 181 GAGTCCGCTTCTATGAGGCGTT-----CTCGCCCCACCGCGCTCCGCGCGCGCATC 234  
Db |||||  
Qy 135 CGGCGCTCCACCGAATCTCTCTTTTACACCTCTCTCGCGCGCGCGCCACGACTACTCC 194  
Db |||||  
Qy 235 CGAGACACTTTTCCCGCGGTTCCAGGCACCGGACTCTCTCCACCGAGGCGCGACGCC 294  
Db |||||  
Qy 195 ATCCGCTCTCTTCTCCCGCGCTTTCAGCGCACCGCGCTCTCTGGACGCTCCGAGGCTCC 254  
Db |||||  
Qy 295 GGGGAGCGGATCTCACTCGCTCTCGACGACTCTCTCGCGGGTTTCAGGCGCCCTGCG 354  
Db |||||  
Qy 255 GG---TCCCAACCTTCTCTGCTGAGACCTCTCTGCGGCTACTCTCAAAACCTTCC 311  
Db |||||  
Qy 355 GTCGACACATCAAGATCGGCGCATCACTGGGCCACCGAGTTCCCGGAGCCCTACATC 414  
Db |||||  
Qy 312 GTCATGGACGTAAAGATCGGCTCCAGAACCTGGCACCTGGGAGACTCCGAGGACTACATC 371  
Db |||||  
Qy 415 GCMAAGTGCCTCGCAAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCGCGGTCTCC 474  
Db |||||  
Qy 372 TGCAAGTGCCTGGAAGAGGACAGAGTCTCTAGCTTGCCTTGGGTTTCAGAAATCTCG 431  
Db |||||  
Qy 475 GGGCTCCGAGTCTGTCGCGCGCGGCGCGTGTGGCGGACGCGCGCGAGGTGAAG 534  
Db |||||  
Qy 432 GGAGTCAAGGACTCTATCTCTCTGGGAACCTTACAGGAAATCTCTCCAGTGTCTATCC 491  
Db |||||  
Qy 535 GCCATGGACACCGCGCGCTCGCGCGGTCTCGCGGCT---ACGTGTCAATCGGTTGCC 591  
Db |||||  
Qy 492 GCCCATGGTGTGACATTGTTCTCAACAAGTTCGTTTCTTAATAATATCAACCATGAT 551  
Db |||||



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2005, 00:32:13 ; Search time 879 Seconds  
(without alignments)  
9491.401 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

| 1:  | /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*  |
|-----|--|
| 2:  | /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*   |
| 3:  | /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*  |
| 4:  | /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*  |
| 5:  | /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*  |
| 6:  | /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.* |
| 7:  | /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*  |
| 8:  | /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*  |
| 9:  | /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.* |
| 10: | /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.* |
| 11: | /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.* |
| 12: | /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*  |
| 13: | /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.* |
| 14: | /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.* |
| 15: | /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.* |
| 16: | /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.* |
| 17: | /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.* |
| 18: | /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.* |
| 19: | /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.* |
| 20: | /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.* |
| 21: | /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.* |
| 22: | /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*  |
| 23: | /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.* |
| 24: | /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*  |
| 25: | /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*  |
| 26: | /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID                | Description        |
|------------|-------|-------------|--------|-------------------|--------------------|
| 1          | 1344  | 100.0       | 1344   | US-10-042-894A-7  | Sequence 7, Appli  |
| 2          | 1307  | 97.2        | 1426   | US-10-042-894A-7  | Sequence 74438, A  |
| 3          | 865.4 | 64.4        | 923    | US-10-042-894A-5  | Sequence 5, Appli  |
| 4          | 844.8 | 62.9        | 1169   | US-10-042-894A-1  | Sequence 1, Appli  |
| 5          | 837.4 | 62.3        | 923    | US-10-042-894A-3  | Sequence 3, Appli  |
| 6          | 740.4 | 55.1        | 3416   | US-10-042-894A-20 | Sequence 20, Appli |
| 7          | 687.4 | 51.1        | 776    | US-10-042-894A-20 | Sequence 74436, A  |

#### ALIGNMENTS

#### RESULT 1

US-10-042-894A-7

; Sequence 7, Application US/10042894A

; Publication No. US2003009011A1

; GENERAL INFORMATION:

; APPLICANT: Shi, Jinrui

; APPLICANT: Beach, Larry

; APPLICANT: Wang, Hongyu

; APPLICANT: Rafalski, Antoni J.

; APPLICANT: Cahoon, Rebecca E.

; TITLE OF INVENTION: No. US2003009011A1e1 Inositol Polyphosphate Kinase

; TITLE OF INVENTION: Genes and Uses Thereof

; FILE REFERENCE: 1286

; CURRENT APPLICATION NUMBER: US/10/042,894A

; CURRENT FILING DATE: 2002-01-09

; PRIOR APPLICATION NUMBER: US 60/261,465

; PRIOR FILING DATE: 2001-01-12

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 1344

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (52)....(921)

; US-10-042-894A-7

Query Match 100.0%; Score 1344; DB 14; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 0;

Sequence 15, Appli  
Sequence 4650, Ap  
Sequence 17, Appli  
Sequence 4649, Ap  
Sequence 18, Appli  
Sequence 10743, A  
Sequence 19, Appli  
Sequence 11, Appli  
Sequence 9, Appli  
Sequence 106467,  
Sequence 13, Appli  
Sequence 340, App  
Sequence 1, Appli  
Sequence 10, Appli  
Sequence 64771, A  
Sequence 57529, A  
Sequence 1, Appli  
Sequence 2855, Ap  
Sequence 1, Appli  
Sequence 86085, A  
Sequence 118038,  
Sequence 5687, Ap  
Sequence 2726, Ap  
Sequence 143130,  
Sequence 829, App  
Sequence 4706, Ap  
Sequence 1, Appli  
Sequence 2917, Ap  
Sequence 33515, A  
Sequence 3065, Ap  
Sequence 1481, Ap  
Sequence 1189, Ap  
Sequence 3940, Ap  
Sequence 50246, A  
Sequence 3431, Ap  
Sequence 15, Appli  
Sequence 1, Appli

51.0 899 14 US-10-042-894A-15  
40.6 1108 19 US-10-043-963-4650  
35.0 643 14 US-10-042-894A-17  
22.9 696 19 US-10-043-963-4649  
22.2 519 14 US-10-042-894A-18  
22.0 295 20 US-10-042-894A-19  
16.4 353 14 US-10-042-894A-11  
13.7 1195 14 US-10-042-894A-9  
13.3 1105 14 US-10-042-894A-9  
13.3 1345 18 US-10-042-894A-13  
11.1 1020 14 US-10-042-894A-13  
7.8 464 9 US-09-770-444-340  
5.3 135638 21 US-10-314-657-1  
5.3 135638 21 US-10-473-193-1  
5.1 1383 18 US-10-417-700A-10  
5.1 1478 19 US-10-437-963-64771  
5.1 2007 19 US-10-437-963-57529  
5.1 88421 9 US-09-976-059-1  
5.1 2541 15 US-10-156-761-2855  
5.1 9025608 15 US-10-156-761-1  
5.0 1095 19 US-10-437-963-86085  
5.0 1347 18 US-10-424-599-118038  
4.9 546 19 US-10-437-963-5687  
4.9 3471 15 US-10-156-761-2726  
4.9 1491 20 US-10-425-115-143130  
4.9 484 10 US-09-770-961-829  
4.8 825 15 US-10-156-761-4706  
4.8 9025608 15 US-10-156-761-1  
4.8 1860 19 US-10-437-963-2917  
4.8 1296 18 US-10-425-114-33515  
4.8 1633 20 US-10-739-930-3065  
4.7 5452 15 US-10-017-161-1481  
4.7 5452 17 US-10-292-798-1189  
4.7 587 15 US-10-156-761-3940  
4.7 1656 19 US-10-437-963-50246  
4.7 1581 15 US-10-156-761-3431  
4.7 1227 21 US-10-399-693-15  
4.7 138203 21 US-10-819-386A-1



|    |      | Matches   | 1344;      | Conservative   | 0;              | Mismatches | 0;        | Indels   | 0;      | Gaps | 0;    |
|----|------|-----------|------------|----------------|-----------------|------------|-----------|----------|---------|------|-------|
| QY | 1    | GCACGAGGT | CAGTCCGTC  | ACCCCTCGCGCCCA | TAGTCCCTTCCCAT  | ACCATGTC   | CCGAC     | 60       |         |      |       |
| Db | 1    | GCACGAGGT | CAGTCCGTC  | ACCCCTCGCGCCCA | TAGTCCCTTCCCAT  | ACCATGTC   | CCGAC     | 60       |         |      |       |
| QY | 61   | CTCCACCCG | CGGACCA    | CAAGTCGCGGCA   | CCCGGCTCCGCA    | CGACGAGT   | GGGCGG    | 120      |         |      |       |
| Db | 61   | CTCCACCCG | CGGAGAC    | CAAGTCGCGGCA   | CCCGGCTCCGCA    | CGACGAGT   | GGGCGG    | 120      |         |      |       |
| QY | 121  | CTCATCGAG | CGGTCCGGCT | CTTCTACAA      | CGCGCTCCAGCGCGG | CGACCGTGGG | GAGAC     | 180      |         |      |       |
| Db | 121  | CTCATCGAG | CGGTCCGGCT | CTTCTACAA      | CGCGCTCCAGCGCGG | CGACCGTGGG | GAGAC     | 180      |         |      |       |
| QY | 181  | GAGGTGCGT | TTCTATGAG  | CGGTTCCTCGCCCA | CGCGCGGTCCGCGC  | CGCATTCG   | AGAC      | 240      |         |      |       |
| Db | 181  | GAGGTGCGT | TTCTATGAG  | CGGTTCCTCGCCCA | CGCGCGGTCCGCGC  | CGCATTCG   | AGAC      | 240      |         |      |       |
| QY | 241  | ACCTTCTTC | CCCGGTTT   | CCACGGCAG      | CGACTCCTCCCA    | CCGAGCGCG  | CGGGGAG   | 300      |         |      |       |
| Db | 241  | ACCTTCTTC | CCCGGTTT   | CCACGGCAG      | CGACTCCTCCCA    | CCGAGCGCG  | CGGGGAG   | 300      |         |      |       |
| QY | 301  | CGGCATCTC | ACCTCGCT   | CGACGAC        | CTCTCGCGGGT     | TTTCAGGCG  | CCCTGCGT  | CGCA     | 360     |      |       |
| Db | 301  | CGGCATCTC | ACCTCGCT   | CGACGAC        | CTCTCGCGGGT     | TTTCAGGCG  | CCCTGCGT  | CGCA     | 360     |      |       |
| QY | 361  | GACATCAAG | ATCGGCGCC  | ATCATCGT       | GGCCACCGAGTT    | TCGCGGAG   | CCCTTACAT | TCGCCAAG | 420     |      |       |
| Db | 361  | GACATCAAG | ATCGGCGCC  | ATCATCGT       | GGCCACCGAGTT    | TCGCGGAG   | CCCTTACAT | TCGCCAAG | 420     |      |       |
| QY | 421  | TGCCTCGCA | AGGACCGCGG | GCACGAG        | GTTCGCTCGGAT    | TCGGGTTC   | CGCGCTC   | 480      |         |      |       |
| Db | 421  | TGCCTCGCA | AGGACCGCGG | GCACGAG        | GTTCGCTCGGAT    | TCGGGTTC   | CGCGCTC   | 480      |         |      |       |
| QY | 481  | CGAGTCTCG | CGGCGCG    | CGGTGTGCGG     | ACGAGCGCG       | CGGAGT     | GAAGC     | CATG     | 540     |      |       |
| Db | 481  | CGAGTCTCG | CGGCGCG    | CGGTGTGCGG     | ACGAGCGCG       | CGGAGT     | GAAGC     | CATG     | 540     |      |       |
| QY | 541  | GACACCGCG | GGTCCGCG   | CGTCTCGG       | CGCTACG         | TCGCTTC    | CGCAC     | AGGG     | 600     |      |       |
| Db | 541  | GACACCGCG | GGTCCGCG   | CGTCTCGG       | CGCTACG         | TCGCTTC    | CGCAC     | AGGG     | 600     |      |       |
| QY | 601  | ATGGACTGT | CGCTCGCG   | CGCGGTGTAC     | GAGGAAG         | GTGGAGT    | CTTGT     | TCACAG   | CTG     | 660  |       |
| Db | 601  | ATGGACTGT | CGCTCGCG   | CGCGGTGTAC     | GAGGAAG         | GTGGAGT    | CTTGT     | TCACAG   | CTG     | 660  |       |
| QY | 661  | CGCGAGCTC | CAAGCGTGT  | TCGAGGAG       | CAGACTCTG       | TCCACTT    | CTACT     | TCGCGT   | CGATT   | 720  |       |
| Db | 661  | CGCGAGCTC | CAAGCGTGT  | TCGAGGAG       | CAGACTCTG       | TCCACTT    | CTACT     | TCGCGT   | CGATT   | 720  |       |
| QY | 721  | CTTCTGGG  | CTATGAT    | CTGCTGAG       | TCGAGCG         | GAGATGGGG  | TGGGTG    | ACGGTG   | 780     |      |       |
| Db | 721  | CTTCTGGG  | CTATGAT    | CTGCTGAG       | TCGAGCG         | GAGATGGGG  | TGGGTG    | ACGGTG   | 780     |      |       |
| QY | 781  | NAGCTGTG  | GAATTTG    | CCCATG         | TGCGGAG         | GTGATGGG   | TGATTC    | ACACAA   | CTTC    | 840  |       |
| Db | 781  | NAGCTGTG  | GAATTTG    | CCCATG         | TGCGGAG         | GTGATGGG   | TGATTC    | ACACAA   | CTTC    | 840  |       |
| QY | 841  | GGCGGGCT  | CTGCTG     | ATCAAGT        | TCGTTTCTG       | ACATTTG    | TCGAGACT  | CTCT     | CATAG   | 900  |       |
| Db | 841  | GGCGGGCT  | CTGCTG     | ATCAAGT        | TCGTTTCTG       | ACATTTG    | TCGAGACT  | CTCT     | CATAG   | 900  |       |
| QY | 901  | CAGGCTTTG | GGTCTCTT   | TAAAGAG        | AGATTC          | TGGCAT     | TTTTCAT   | TAAACAA  | AGCCCTA | 960  |       |
| Db | 901  | CAGGCTTTG | GGTCTCTT   | TAAAGAG        | AGATTC          | TGGCAT     | TTTTCAT   | TAAACAA  | AGCCCTA | 960  |       |
| QY | 961  | CAAGTTTCT | GTGGA      | AAAGAG         | CGCTCCGAG       | TGTGCTGG   | TGGAG     | TC       | TGAG    | 1020 |       |
| Db | 961  | CAAGTTTCT | GTGGA      | AAAGAG         | CGCTCCGAG       | TGTGCTGG   | TGGAG     | TC       | TGAG    | 1020 |       |
| QY | 1021 | GTGCTGGG  | CCCACT     | TGGTTG         | CCCTTGC         | CTTTCG     | CTTGA     | AAACAT   | ACGG    | CAAC | CTGCT |
| Db | 1021 | GTGCTGGG  | CCCACT     | TGGTTG         | CCCTTGC         | CTTTCG     | CTTGA     | AAACAT   | ACGG    | CAAC | CTGCT |

|  |      |  |      |
|--|------|--|------|
| QY   | 1081 | CTTTTTTTCGCAACCCCTTACTTCCGAAGAAACTTTTTTTTCCCACTTTGGGGGTTTCGA   | 1140 |
| Db   | 1081 | CTTTTTTTCGCAACCCCTTACTTCCGAAGAAACTTTTTTTTCCCACTTTGGGGGTTTCGA   | 1140 |
| QY   | 1141 | TTACGTTGGATCTGGTTTGTGCCACTCGGTACAGAGTTGTAAAGCATGGAGGAGGCGGTGT  | 1200 |
| Db   | 1141 | TTACGTTGGATCTGGTTTGTGCCACTCGGTACAGAGTTGTAAAGCATGGAGGAGGCGGTGT  | 1200 |
| QY   | 1201 | TGATCCGGCAACTGTGTCACTTTTCGGCTGCGCTGCGCTTTCTGTCATGGCTTTTGGCCTGC | 1260 |
| Db   | 1201 | TGATCCGGCAACTGTGTCACTTTTCGGCTGCGCTGCGCTTTCTGTCATGGCTTTTGGCCTGC | 1260 |
| QY   | 1261 | TGCGATCCGATGTGTACTTGGAGATCGTAGTGATGGAAGCTCTTACTCTCCAAACGAATCCG | 1320 |
| Db   | 1261 | TGCGATCCGATGTGTACTTGGAGATCGTAGTGATGGAAGCTCTTACTCTCCAAACGAATCCG | 1320 |
| QY   | 1321 | TCCGATAAAAAAAAAAAAAAAAAAAAAA                                   | 1344 |
| Db   | 1321 | TCCGATAAAAAAAAAAAAAAAAAAAAAA                                   | 1344 |
| RESULT 2   |      |  |      |
| US-10-425-115-74438  |      |  |      |
| ; Sequence 74438, Application US/10425115  |      |  |      |
| ; Publication No. US20040214272A1  |      |  |      |
| ; GENERAL INFORMATION:   |      |  |      |
| ; APPLICANT: La Rosa, Thomas J.  |      |  |      |
| ; APPLICANT: Kovalic, David K.   |      |  |      |
| ; APPLICANT: Zhou, Yihua   |      |  |      |
| ; APPLICANT: Cao, Yongwei  |      |  |      |
| ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With |      |  |      |
| ; FILE REFERENCE: 38-21(5322)B   |      |  |      |
| ; CURRENT APPLICATION NUMBER: US/10/425,115                                      |      |  |      |
| ; CURRENT FILING DATE: 2003-04-28  |      |  |      |
| ; NUMBER OF SEQ ID NOS: 369326   |      |  |      |
| ; SEQ ID NO 74438  |      |  |      |
| ; LENGTH: 1426   |      |  |      |
| ; TYPE: DNA  |      |  |      |
| ; ORGANISM: Zea mays   |      |  |      |
| ; FEATURE:   |      |  |      |
| ; OTHER INFORMATION: Clone ID: MRT4577_167886C.1                                 |      |  |      |
| US-10-425-115-74438  |      |  |      |
| Query Match 97.2%; Score 1307; DB 20; Length 1426;                               |      |  |      |
| Best Local Similarity 99.5%; Pred. No. 0;  |      |  |      |
| Matches 1321; Conservative 0; Mismatches 5; Indels 1; Gaps 1;                    |      |  |      |
| QY   | 8    | GTCACTCCGTCAACCCCTCGCGCCATAGTCCCTTCCCATACCATGTCCGACCTCCACC     | 67   |
| Db   | 76   | GTCACTCCGTCAACCCCTCGCGCCATAGTCCCTTCCCATACCATGTCCGACCTCCACC     | 135  |
| QY   | 68   | CGCGGAGACCAAGTCGCGCGGCCAACCGCGCTCCGCGCAGCAAGCTGGGCGCGCTCATCG   | 127  |
| Db   | 136  | CGCGGAGACCAAGTCGCGCGGCCAACCGCGCTCCGCGCAGCAAGCTGGGCGCGCTCATCG   | 195  |
| QY   | 128  | ACGGCTCCGGCTCTTCTACAGCCGCTCCAGCCCGCGACCGTGGGGACACGAGGTGCG      | 187  |
| Db   | 196  | ACGGCTCCGGCTCTTCTACAGCCGCTCCAGCCCGCGACCGTGGGGACACGAGGTGCG      | 255  |
| QY   | 188  | CCTTCTATGAGGCGTTCTTCGCCACCGCGCCCTCCCGGCCCATCCGAGACACCTTCT      | 247  |
| Db   | 256  | CCTTCTATGAGGCGTTCTTCGCCACCGCGCCCTCCCGGCCCATCCGAGACACCTTCT      | 315  |
| QY   | 248  | TCCCGGTTTCCACGCAACGCACTCCTCCCAACGAGGCGCAGCCGGGAGCCCGCATC       | 307  |
| Db   | 316  | TCCCGGTTTCCACGCAACGCACTCCTCCCAACGAGGCGCAGCCGGGAGCCCGCATC       | 375  |
| QY   | 308  | CTCACCTCGTCTCGACGACCTCTCGCGGGGTTTCAGGGCGCCTGCTCGCAGACATCA      | 367  |
| Db   | 376  | CTCACCTCGTCTCGACGACCTCTCGCGGGGTTTCAGGGCGCCTGCTCGCAGACATCA      | 435  |

```

368 AGATCGGCGCATACGTCGCCACCGAGTTCCGCGGAGCCCTACATCGCAAGTGCCTCG 427
369 |||||
370 |||||
371 |||||
372 |||||
373 |||||
374 |||||
375 |||||
376 |||||
377 |||||
378 |||||
379 |||||
380 |||||
381 |||||
382 |||||
383 |||||
384 |||||
385 |||||
386 |||||
387 |||||
388 |||||
389 |||||
390 |||||
391 |||||
392 |||||
393 |||||
394 |||||
395 |||||
396 |||||
397 |||||
398 |||||
399 |||||
400 |||||
401 |||||
402 |||||
403 |||||
404 |||||
405 |||||
406 |||||
407 |||||
408 |||||
409 |||||
410 |||||
411 |||||
412 |||||
413 |||||
414 |||||
415 |||||
416 |||||
417 |||||
418 |||||
419 |||||
420 |||||
421 |||||
422 |||||
423 |||||
424 |||||
425 |||||
426 |||||
427 |||||
428 CCAAGAACCGCGGACCAAGAGCGTTCTGTCGGATTCCGCGTCTCCGAGTCCGAGTCC 487
429 |||||
430 |||||
431 |||||
432 |||||
433 |||||
434 |||||
435 |||||
436 |||||
437 |||||
438 |||||
439 |||||
440 |||||
441 |||||
442 |||||
443 |||||
444 |||||
445 |||||
446 |||||
447 |||||
448 |||||
449 |||||
450 |||||
451 |||||
452 |||||
453 |||||
454 |||||
455 |||||
456 |||||
457 |||||
458 |||||
459 |||||
460 |||||
461 |||||
462 |||||
463 |||||
464 |||||
465 |||||
466 |||||
467 |||||
468 |||||
469 |||||
470 |||||
471 |||||
472 |||||
473 |||||
474 |||||
475 |||||
476 |||||
477 |||||
478 |||||
479 |||||
480 |||||
481 |||||
482 |||||
483 |||||
484 |||||
485 |||||
486 |||||
487 |||||
488 |||||
489 |||||
490 |||||
491 |||||
492 |||||
493 |||||
494 |||||
495 |||||
496 |||||
497 |||||
498 |||||
499 |||||
500 |||||
501 |||||
502 |||||
503 |||||
504 |||||
505 |||||
506 |||||
507 |||||
508 |||||
509 |||||
510 |||||
511 |||||
512 |||||
513 |||||
514 |||||
515 |||||
516 |||||
517 |||||
518 |||||
519 |||||
520 |||||
521 |||||
522 |||||
523 |||||
524 |||||
525 |||||
526 |||||
527 |||||
528 |||||
529 |||||
530 |||||
531 |||||
532 |||||
533 |||||
534 |||||
535 |||||
536 |||||
537 |||||
538 |||||
539 |||||
540 |||||
541 |||||
542 |||||
543 |||||
544 |||||
545 |||||
546 |||||
547 |||||
548 |||||
549 |||||
550 |||||
551 |||||
552 |||||
553 |||||
554 |||||
555 |||||
556 |||||
557 |||||
558 |||||
559 |||||
560 |||||
561 |||||
562 |||||
563 |||||
564 |||||
565 |||||
566 |||||
567 |||||
568 |||||
569 |||||
570 |||||
571 |||||
572 |||||
573 |||||
574 |||||
575 |||||
576 |||||
577 |||||
578 |||||
579 |||||
580 |||||
581 |||||
582 |||||
583 |||||
584 |||||
585 |||||
586 |||||
587 |||||
588 |||||
589 |||||
590 |||||
591 |||||
592 |||||
593 |||||
594 |||||
595 |||||
596 |||||
597 |||||
598 |||||
599 |||||
600 |||||
601 |||||
602 |||||
603 |||||
604 |||||
605 |||||
606 |||||
607 |||||
608 |||||
609 |||||
610 |||||
611 |||||
612 |||||
613 |||||
614 |||||
615 |||||
616 |||||
617 |||||
618 |||||
619 |||||
620 |||||
621 |||||
622 |||||
623 |||||
624 |||||
625 |||||
626 |||||
627 |||||
628 |||||
629 |||||
630 |||||
631 |||||
632 |||||
633 |||||
634 |||||
635 |||||
636 |||||
637 |||||
638 |||||
639 |||||
640 |||||
641 |||||
642 |||||
643 |||||
644 |||||
645 |||||
646 |||||
647 |||||
648 |||||
649 |||||
650 |||||
651 |||||
652 |||||
653 |||||
654 |||||
655 |||||
656 |||||
657 |||||
658 |||||
659 |||||
660 |||||
661 |||||
662 |||||
663 |||||
664 |||||
665 |||||
666 |||||
667 |||||
668 |||||
669 |||||
670 |||||
671 |||||
672 |||||
673 |||||
674 |||||
675 |||||
676 |||||
677 |||||
678 |||||
679 |||||
680 |||||
681 |||||
682 |||||
683 |||||
684 |||||
685 |||||
686 |||||
687 |||||
688 |||||
689 |||||
690 |||||
691 |||||
692 |||||
693 |||||
694 |||||
695 |||||
696 |||||
697 |||||
698 |||||
699 |||||
700 |||||
701 |||||
702 |||||
703 |||||
704 |||||
705 |||||
706 |||||
707 |||||
708 |||||
709 |||||
710 |||||
711 |||||
712 |||||
713 |||||
714 |||||
715 |||||
716 |||||
717 |||||
718 |||||
719 |||||
720 |||||
721 |||||
722 |||||
723 |||||
724 |||||
725 |||||
726 |||||
727 |||||
728 |||||
729 |||||
730 |||||
731 |||||
732 |||||
733 |||||
734 |||||
735 |||||
736 |||||
737 |||||
738 |||||
739 |||||
740 |||||
741 |||||
742 |||||
743 |||||
744 |||||
745 |||||
746 |||||
747 |||||
748 |||||
749 |||||
750 |||||
751 |||||
752 |||||
753 |||||
754 |||||
755 |||||
756 |||||
757 |||||
758 |||||
759 |||||
760 |||||
761 |||||
762 |||||
763 |||||
764 |||||
765 |||||
766 |||||
767 |||||
768 |||||
769 |||||
770 |||||
771 |||||
772 |||||
773 |||||
774 |||||
775 |||||
776 |||||
777 |||||
778 |||||
779 |||||
780 |||||
781 |||||
782 |||||
783 |||||
784 |||||
785 |||||
786 |||||
787 |||||
788 |||||
789 |||||
790 |||||
791 |||||
792 |||||
793 |||||
794 |||||
795 |||||
796 |||||
797 |||||
798 |||||
799 |||||
800 |||||
801 |||||
802 |||||
803 |||||
804 |||||
805 |||||
806 |||||
807 |||||
808 |||||
809 |||||
810 |||||
811 |||||
812 |||||
813 |||||
814 |||||
815 |||||
816 |||||
817 |||||
818 |||||
819 |||||
820 |||||
821 |||||
822 |||||
823 |||||
824 |||||
825 |||||
826 |||||
827 |||||
828 |||||
829 |||||
830 |||||
831 |||||
832 |||||
833 |||||
834 |||||
835 |||||
836 |||||
837 |||||
838 |||||
839 |||||
840 |||||
841 |||||
842 |||||
843 |||||
844 |||||
845 |||||
846 |||||
847 |||||
848 |||||
849 |||||
850 |||||
851 |||||
852 |||||
853 |||||
854 |||||
855 |||||
856 |||||
857 |||||
858 |||||
859 |||||
860 |||||
861 |||||
862 |||||
863 |||||
864 |||||
865 |||||
866 |||||
867 |||||
868 |||||
869 |||||
870 |||||
871 |||||
872 |||||
873 |||||
874 |||||
875 |||||
876 |||||
877 |||||
878 |||||
879 |||||
880 |||||
881 |||||
882 |||||
883 |||||
884 |||||
885 |||||
886 |||||
887 |||||
888 |||||
889 |||||
890 |||||
891 |||||
892 |||||
893 |||||
894 |||||
895 |||||
896 |||||
897 |||||
898 |||||
899 |||||
900 |||||
901 |||||
902 |||||
903 |||||
904 |||||
905 |||||
906 |||||
907 |||||
908 |||||
909 |||||
910 |||||
911 |||||
912 |||||
913 |||||
914 |||||
915 |||||
916 |||||
917 |||||
918 |||||
919 |||||
920 |||||
921 |||||
922 |||||
923 |||||
924 |||||
925 |||||
926 |||||
927 |||||
928 |||||
929 |||||
930 |||||
931 |||||
932 |||||
933 |||||
934 |||||
935 |||||
936 |||||
937 |||||
938 |||||
939 |||||
940 |||||
941 |||||
942 |||||
943 |||||
944 |||||
945 |||||
946 |||||
947 |||||
948 |||||
949 |||||
950 |||||
951 |||||
952 |||||
953 |||||
954 |||||
955 |||||
956 |||||
957 |||||
958 |||||
959 |||||
960 |||||
961 |||||
962 |||||
963 |||||
964 |||||
965 |||||
966 |||||
967 |||||
968 |||||
969 |||||
970 |||||
971 |||||
972 |||||
973 |||||
974 |||||
975 |||||
976 |||||
977 |||||
978 |||||
979 |||||
980 |||||
981 |||||
982 |||||
983 |||||
984 |||||
985 |||||
986 |||||
987 |||||
988 |||||
989 |||||
990 |||||
991 |||||
992 |||||
993 |||||
994 |||||
995 |||||
996 |||||
997 |||||
998 |||||
999 |||||
1000 |||||

```

RESULT 3  
 US-10-042-894A-5  
 ; Sequence 5, Application US/10042894A  
 ; Publication No. US20030009011A1

```

; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(922)
; US-10-042-894A-5

Query Match      64.4%; Score 865.4; DB 14; Length 923;
Best Local Similarity 98.8%; Pred. No. 2.8e-236;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 42 TCCCATACCATGTTCGACCTCACCCGCGGAGACCAAGTGCCTCGCGGCGCACCGCGCTC 101
DB 43 TCCCATACCATGTTCGACCTCACCCGCGGAGACCAAGTGCCTCGCGGCGCACCGCGCTC 102
QY 102 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 161
DB 103 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 162
QY 162 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 221
DB 163 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 222
QY 222 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 281
DB 223 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 282
QY 282 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 341
DB 283 CGCGAGACGTGGCGCGCGCTCATCGACGGCTCCGGCCCTTCTACAAGCGCTCCAGGC 342
QY 342 TCAGGCGCGCTCGTCGACAGATCAAGATCGCGCGCTTCTACAAGCGCTCCAGGC 401
DB 343 TCAGGCGCGCTCGTCGACAGATCAAGATCGCGCGCTTCTACAAGCGCTCCAGGC 402
QY 402 CGAGCGCTTACATCGCGCAAGTGCCTCGCGAGACCGCGGGAACACAGAGGTTCTGTCGG 461
DB 403 CGAGCGCTTACATCGCGCAAGTGCCTCGCGAGACCGCGGGAACACAGAGGTTCTGTCGG 462
QY 462 ATTTCGCGCTCCGCGCTCCGAGTCCGCGCGCGCGCGCGCTTCTACAAGCGCTCCAGGC 521
DB 463 ATTTCGCGCTCCGCGCTCCGAGTCCGCGCGCGCGCGCGCTTCTACAAGCGCTCCAGGC 522
QY 522 CGCGAGGTGAAGGCGCATCGACACCGCGCGCGCTCGCGCGCGCTTCTACAAGCGCTCCAGGC 581
DB 523 CGCGAGGTGAAGGCGCATCGACACCGCGCGCGCTCGCGCGCGCTTCTACAAGCGCTCCAGGC 582
QY 582 ATCCGTTGCCGACGAGGGGATGGAATGCTGCTCGCGCGCGCTTCTACAAGCGCTCCAGGC 641
DB 583 ATCCGTTGCCGACGAGGGGATGGAATGCTGCTCGCGCGCGCTTCTACAAGCGCTCCAGGC 642
QY 642 TGGAGTCTTGTTCACAGCTCGCGAGCTCAAGCGGTGTTTCGAGGACGAGACTCTGTTC 701
DB 643 TGGAGTCTTGTTCACAGCTCGCGAGCTCAAGCGGTGTTTCGAGGACGAGACTCTGTTC 702
QY 702 CTCTACTCGGGCTCGATTTCTTCTGGGCTATGATGCTGCTGACAGTCCGAGCGGAG 761

```

```
Db 703 CTTCTACTCGGCGTCGATTCTTCTGGCTATATGCTGTGAGTCGACGAGCGGAGG 762
QY 762 TGGGGGTGGGTGACGGTGAAGCTGGTGAAGCTTTGCCCCATGTGGCCGAGGGGTGATGGGGT 821
Db 763 TGGGGGTGGGTGACGGTGAAGCTGGTGAAGCTTTGCCCCATGTGGCCGAGGGGTGATGGGGT 822
QY 822 GATTGACCAACTTCTTGGGGGGCTCTGCTGGCTGATCAAGTTTGGTTTCTTGACATTGT 881
Db 823 GATTGACCAACTTCTTGGGGGGCTCTGCTGGCTGATCAAGTTTGGTTTCTTGACATTGT 882
QY 882 TCCGAGACTCTCTATACAGGCTTTTGGGTCTTCTTAAG 922
Db 883 TCCAGAGACTCTCTAGACGAGCCTTTGGGTCTTCTTAAG 923

RESULT 4
US-10-042-894A-1
; Sequence 1, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(806)
US-10-042-894A-1

Query Match 62.9%; Score 844.8; DB 14; Length 1169;
Best Local Similarity 97.2%; Pred. No. 2.2e-230;
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 42 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCCTC 101
Db 74 TCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCCTC 133
QY 102 CGCCAGCAAGCTGGGCGGCTCATCGAGGCTCCGCGCTTCTTCTAAGCGCTCCAGGC 161
Db 134 CGCCAGCAAGCTGGGCGGCTCATCGAGGCTCCGCGCTTCTTCTAAGCGCTCCAGGC 193
QY 162 CGCGAGCGTGGGAGCAGAGTGCCTTCTATAGGGGTTCTCGGCCACCGCGCGGT 221
Db 194 CGCGAGCGTGGGAGCAGAGTGCCTTCTATAGGGGTTCTCGGCCACCGCGCGGT 253
QY 222 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCAGCGCAGCGACTCTTCCCCAC 281
Db 254 CCGGCGCGCATCCGAGACACCTTCTTCCCGGTTTCCAGCGCAGCGACTCTTCCCCAC 313
QY 282 CGAGGCGCAGCGCGGAGCCGATCCTCTCAGCAGCTCTCTCGCGGGTT 341
Db 314 CGAGGCGCAGCGCGGAGCCGATCCTCGCACTCTCTCGAGCAGCTCTCTCGCGGGTT 373
QY 342 TCAGGCGCCTCGTCCGAGACATCAAGATCGCGGCATCAGTGGCCACCGAGTTCCGC 401
Db 374 TGAGGCGCCTCGTCCGAGACATCAAGATCGCGGCATCAGTGGCCACCGAGTTCCGC 433
QY 402 GGAGCGCTACATCGCCAAAGTGTCTCGCCAAAGACCGCGGACACAGAGGTTCTGTCTGG 461
```

```
Db 434 GGAGCGCTACATCGCAAGTACCTCCACAGGACCGCGGACCAAGAGGTTCTGTCTCG 493
QY 462 ATTCCGCTCTCCGCGTCCGAGTCTGTCGCCCCCGAGGCGCGGTGTGCGCAGCGAGCG 521
Db 494 ATTCCGCTCT--TGCGTCCGAGTCTGTCGCCCCCGAGGCGCGGTGTGCGCAGCGAGCG 551
QY 522 CCCGAGGTGAAGCCATGGACACCGCGGCGTCCGCGCGTCTCCGCGCTACGTGTC 581
Db 552 CCCGAGGTGAAGGTATGGACACCGTCCGCGGTCCGCGCGTCTCCGCGCTACGTGTC 611
QY 582 ATCCG--TTGCCGACGAGGGGATGGACTGTGCGCTCCGCGCGGTGTACGAGGAAAAG 640
Db 612 ATCCGCTTCCGACGAGGGGATGGACTGTGCGCTCCGCGCGGTGTACGAGGAAAAG 671
QY 641 GTGGAGTCTTGTCAAGCTGCGGAGCTCAAGCGCTGTGTCGAGGAGAGACTCTGTTC 700
Db 672 GTGGAGTCTTGTCAAGCTGCGGAGCTCAAGGCAATGGTTGGAGGAGCAGACTCTGTTC 731
QY 701 ACTTCTACTCGGCGTCTGATTCTTCTGGGCTATGATGCTGCTGAGTCCGAGCGGAG 760
Db 732 ACTTCTACTCGGCGTCTGATTCTTCTGGGCTATGATGCTGCTGAGTCCGAGCGGAG 791
QY 761 ATGGGGTGGGTGACGCTGAAGCTGTGGGACTTTGCCCATGTGGCCGAGGATGGGG 820
Db 792 GTGGGGTGGGTGACGCTGAAGCTGTGGGACTTTGCCCATGTGGCCGAGGATGGGG 851
QY 821 TGATTGACACAACTTCTGCGGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTG 880
Db 852 TGATTGACACAACTTCTGCGGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTG 911
QY 881 TTCCGAGACTCTCTATAGCAGCGCTTTGGGTCTTCTTAAAGAGAGGATCCCTGSCA-TTT 939
Db 912 TTCCAGAGACTCTCTTAGACGAGCGCTTTGGGTCTTCTTAAAGAGAGGATCCCTGACATTTT 971
QY 940 CGATTGTATAACAAAG 955
Db 972 TGATTGTATACAAAG 987

RESULT 5
US-10-042-894A-3
; Sequence 3, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(736)
US-10-042-894A-3

Query Match 62.3%; Score 837.4; DB 14; Length 923;
Best Local Similarity 97.5%; Pred. No. 2.7e-228;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 42 TCCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCCACCGCGCCTC 101
```

|  |   |   |   |     |
|--|---|---|---|-----|
|  | D | b | T C C C A T A C C A T G C C G G A C C T C C A C C G C C G G A G A C C A A G T G C G C G G T C A C C G G C G C T C     | 102 |
|  | Q | y | G C C A G C A A G C T G G G C C C G C T C A T C G A C G G C T C G G C C T T C T T C T A C A A G C C G C T C C A G C G | 161 |
|  | D | b | G C C A G C A A G C C G G C C G C T C A T C G A C G G C T C C G G C T T C T T C A C A G C O G C T C C A G C G         | 162 |
|  | Q | y | C G G C A C C G T G G G A G A C A G A G T G C C T T C T A T A G A G C G T T C T C G C C C A C G C C G C G T           | 221 |
|  | D | b | C G G C A C C G T G G G A G C A G A G T G C T T C T A T A G A G C G T T C T C G C C C A C G C C G C G T               | 222 |
|  | Q | y | C C G G C C G C A T C C G A G A C A C T T C T T C C C C G G T T C A C G S C A C G C G A C T C T C C C A C             | 281 |
|  | D | b | C C G G C C G C A T C C G A G A C A C T T C T T C C C C G G T T C C A C G S C A C G C G A C T C T C C C A C           | 282 |
|  | Q | y | C G A G C G A C G C C G G G A G C C G A C T C T C A C C T G C T C T C G A C G A C C T C C T C G C G G G T T           | 341 |
|  | D | b | C G A G C G A C G C C G G G A G C C G A C T C G C A C C T G T C C T C G A C G A C C T C C T C G G G A T T             | 342 |
|  | Q | y | T C A G G C C C T G C T G C G A G A C A T C A A G A T C G G G C A T C A C G T G G C C A C C G A G T T C G C C         | 401 |
|  | D | b | T G A G C G C C T C G T C G C A G A C A T C A A G A T C G G C C C A T C A G T G T G C C A C C G A G T T C G C C       | 402 |
|  | Q | y | G G A G C C T A C A T C G C C A A G T G C T C G C C A A G A C C G C G G A C C A C G A G A G T T C T G C T C G G       | 461 |
|  | D | b | G G A G C C T A C A T C G C C A A G T G C T C G C C A A G A C C G C G G A C C A C G A G A G T T C T G C T C G G       | 462 |
|  | Q | y | A T T C G C G T C T C C G G C G T C C G A G T G T C G G C C C C G A G G C G C G T G T G C G G A C G G A G C G         | 521 |
|  | D | b | A T T C G C G T C T C C G G C G T C C G A G T G T C G G C C C C G A G G C G C G T G T G C G G A C G G A G C G         | 522 |
|  | Q | y | C C C G A G T G A A G G C C A T G A C A C C G C C G G C G T C C G C C G C G T C C C G C G C G T A C G T G T C         | 581 |
|  | D | b | C C C G A G T G A A G G C C A T G A C A C C G C C G G C G T C C G C C G C G T C C C G C G C G T A C G T G T C         | 582 |
|  | Q | y | A T C C G - T T G C C G A C G A G G G A T G G A C T G C C C T C G C C G C G G G T G T A C G G A G G A A A G           | 640 |
|  | D | b | A T C C G C T T T G C C G A C A G A G G G A T G G A C T G C G C C T C G C C G C G G G T G T A C G G A G G A A A G     | 642 |
|  | Q | y | G T G G A G T C T T G T C A C A G C T C G C G A G C T C A A G C G T G T T C G A G G A C G A G A C T C T G T T C C     | 700 |
|  | D | b | G T G G A G T C T T G T C A C T G T C G C G A G C T C A A G C G T G T T C G A G G A C G A G A C T C T G T T C C       | 702 |
|  | Q | y | A C T T C T A C T C G C G T C G A T T C T T T G G G C T A T E A T G C T C T G C A G T C G C A G A G C G G A G         | 760 |
|  | D | b | A C T T C T A C T C G C G T C G A T T C T T T G G G C T A T E A T G C T C T G C A G T C G C A G A G C G G A G         | 762 |
|  | Q | y | A T G G G G T G G G T G A C G T G A A C T G G T G G A C T T T G C C C A T G T G G C C G A G G T G A T G G G G         | 820 |
|  | D | b | G T G G G G T G G G T A A C A G T G A A G C T G T G G A C T T T T G C C C A T G T G G C C G A G G G T A T G G G G     | 822 |
|  | Q | y | T G A T T G A C C A A A C T T C C T G G G C G G C T C T G C T G C T G A T C A A G T T C G T T T C T G A C A T T G     | 880 |
|  | D | b | T G A T T G A C C A A A C T T C C T G G G C G G C C T C T G C T A G C T G A T C A A G T T C G T T C T G A C A T T G   | 882 |
|  | Q | y | T T C G G A G A C T C C T C A T A C G A C G C C T T T G G G T C C T T C T T A A                                       | 921 |
|  | D | b | T T C C A G A G A C T C C T C A G A C G C A G C C T T T G G G T C C T T C T T A A                                     | 923 |

## RESULT 6

```

US-10-042-894A-20
; Sequence 20, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalecki, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286

```

```

; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SEQUENCE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 3416
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72) ... (407)
; US-10-042-894A-20

```

| Query Match           | 55.1%          | Score 740.4                              | DB 14                            | Length 3416 |
|-----------------------|----------------|--|----------------------------------|-------------|
| Best Local Similarity | 90.2%          | Pred. No. 1.7e-200                       |                                  |             |
| Matches 862           | Conservative 0 | Mismatches 31                            | Indels 63                        | Gaps 4      |
| QY                    | 2              | CACGAGGTGAGTCGCTCACCCCTCGGCGCCATAGTCCTTC | CCCATAG-----TCCCATATACATGCCCGACC | 61          |
| DB                    | 28             | CACATCGCACCGGTACCCCTTGCTCCCATAG-----     |                                  | 81          |
| QY                    | 62             | TCCACCCGCGGAGCACAAAGTCGCGCGCCACCGCGCTTC  | CGCAGCAAGCTGGGCCCGC              | 121         |
| DB                    | 82             | TCCACCCGCGGAGCACAAAGTCGCGGTCAACCGGCTC    | CGCAGCAAGCTGGGCCCGC              | 141         |
| QY                    | 122            | TCATCGAGCGCTCGGCGCTTCTACAGCGGTTCAGGCGG   | CGGAGCCGTGGGAGACAG               | 181         |
| DB                    | 142            | TCATCGAGCACTCTGGCGCTTCTACAAAGCGCTTCA     | AGCGCGGACCGCTGGGAGACG            | 201         |
| QY                    | 182            | AGTTCGCTTCTATAGGCGGTTCGCGCCACGCGCGCTC    | CGGCGCGCATCCGAGACA               | 241         |
| DB                    | 202            | AGTTCGCTTCTATAGGCGGTTCGCGCCACGCGCGCTC    | CGGCGCGCATCCGAGACA               | 261         |
| QY                    | 242            | CTTCTTTCCCGGTTCCAGCGACGCGACTCTCTCCAC     | CGAGGCGCAGCCGGGAGC               | 301         |
| DB                    | 262            | CTTCTTTCCCGGTTCCAGCGAGCGACTCTCTCCAC      | CGAGGCGCAGCCGGGAGC               | 321         |
| QY                    | 302            | CGCATCTACCTGCTCTCGACGACTCTCTCGGGGGTTT    | CAGGCGCGCTTGCTCGCAG              | 361         |
| DB                    | 322            | CGCATCTCGACCTCTGCTCGACGACTCTCTCGGGGGT    | TGAGGCGCGCTTGCTCGCAG             | 381         |
| QY                    | 362            | ACATCAAGATCGGCGCATACAGCTGGGCCACCGAGT     | TCGCGGAGCCCTACATCGCCAAAT         | 421         |
| DB                    | 382            | ACATCAAGATCGGTGCCATCAGT-----             |                                  | 405         |
| QY                    | 422            | GCCTCGCCAAAGACCGCGGACCAACGAGCGTTCTGT     | TCGATTCCGCTTCGCGCGTCC            | 481         |
| DB                    | 406            | -----GACCACGAGCGTTCTGCTCGATTCCGCGT       | CTCGCGCGTCC                      | 446         |
| QY                    | 482            | GAGTCGTGCGCCGCGAGGGCGGTGTGGCGGACGAG      | CGCGCCCGGAGGTGAGGCCATGG          | 541         |
| DB                    | 447            | GAGTCGTGCGCCGCGAGGGCGGTGTGGCGGACGAG      | CGCGCCCGGAGGTGAGGCCATGG          | 506         |
| QY                    | 542            | ACACGCGCGCTCGCGCGGTGCTCGCGCGCTACGTGT     | CATCGCTTCGCGAGCGGGG              | 600         |
| DB                    | 507            | ACATTGTGCGGCTCCGCGCGGTGCTCGCGCGCTAC      | GTGCTCATCGCTTCGCGAGGGG           | 566         |
| QY                    | 601            | ATGACATGTGCGCTCGCGCGCGGTGTACGAGGAAAG     | AGGTGAGTCTTGTGTACAGCTG           | 660         |
| DB                    | 567            | ATGACATGTGCGCTCGCGCGCGGTGTACGAGGAAAG     | GTGAGTCTTGTGTACAGCTG             | 626         |
| QY                    | 661            | CGCAGCTCAAGGCGTGGTTCGAGGACAGACTGTTC      | CACTCTACTCGGCGTCCATT             | 720         |
| DB                    | 627            | CGCAGCTCAAGGCGTGGTTCGAGGCGACACTGTTC      | CACTCTACTCGGCGTCCATT             | 686         |
| QY                    | 721            | CTTCTGGGCTATGATGCTGCTGACATCGCAGCAGG      | CGGAGATGGGGGTGGGTCACGGTG         | 780         |
| DB                    | 687            | CTTCTGGGCTATGATGCTGCTGACATCGCAGCAGG      | CGGAGATGGGGGTGGGTCACGGTG         | 746         |
| QY                    | 781            | AAGCTGCTGGACTTTGCCCATGTGGCCCGAGGGTGA     | TGAGGGGTGATGACCAACAATCTCTG       | 840         |

Db 747 AAGCTGGTGGACTTTGCCCATGTGGCGAGGGTGATGGGGTGATTGACCAACAATTCCTG 806  
Qy 841 GGGGGGCTCTGCTCGCTGATCAAGTTTCGTTTCGACATTTGTTCCGGAGACTCTCTATAG 900  
Db 807 GGGGGGCTCTGCTAGCTGATCAAGTTTGTTCGACATTTGTTCCAGAGACTCTCTCAGACG 866  
Qy 901 CAGCCTTTGGGTCTCTTTAAAGAGAGAGTCCCTGGCA-TTTTCGATTTGATAACAAG 955  
Db 867 CAGCCCTTGGGTCTCTTTAAAGAGAGAGTCCCTGGCATTTTCGATTTGATAACAAG 922

RESULT 7  
US-10-425-115-74436  
; Sequence 74436, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 74436  
; LENGTH: 776  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(776)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_167884C.1  
US-10-425-115-74436

Query Match 51.1%; Score 687.4; DB 20; Length 776;  
Best Local Similarity 99.7%; Pred. No. 1.6e-185;  
Matches 699; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 566 TCCGGGCTACGTTGTCATCGTTGCCGACGAGGGATGGACTGTGGCTCGCGGGCGG 625  
Db 77 TCCGGGCTACGTTGTCATCGTTGCCGACGAGGGATGGACTGTGGCTCGCGGGCGG 136  
Qy 626 TGTACGGAGGAAAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGGGTGGTTTCGAG 685  
Db 137 TGTACGGAGGAAAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGGGTGGTTTCGAG 196  
Qy 686 AGCAGACTCTGTCCACTTCTACTCGGCGTCGATTCCTTCTGGGCTATGATGCTGTCGAG 745  
Db 197 AGCAGACTCTGTTCACCTTCTACTCGGCGTCGATTCCTTCTGGGCTATGATGCTGTCGAG 256  
Qy 746 TCGCAGCAGCGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGG 805  
Db 257 TCGCAGCAGCGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGG 316  
Qy 806 CCGAGGGTGATGGGGTGATTTGACCAACAATTCCTGGGCGGGCTCTGCTCGCTGATCAAGT 865  
Db 317 CCGAGGGTGATGGGGTGATTTGACCAACAATTCCTGGGCGGGCTCTGCTCGCTGATCAAGT 376  
Qy 866 TCGTTTCTACATTTGTTCCGGAGACTCTCATACGAGGCTTTGGTTCCTTCTTAAGAGA 925  
Db 377 TCGTTTCTACATTTGTTCCGGAGACTCTCATACGAGGCTTTGGGTCCTTCTTAAGAGA 436  
Qy 926 GGATCCTGGCATTTTCGATTTGATAACAAGCCCTACAAGTTTGTCTGGAAGAAAGAGCG 985  
Db 437 GGATCCTGGCATTTTCGATTTGATAACAAGCCCTACAAGTTTGTCTGGAAGAAAGAGCG 496  
Qy 986 CTTCCGAGTTGTCTGGGTGTGGAGATCTGAGACGGTCTGCGGCCCACTTGGTTGCTTGG 1045  
Db 497 CTTCCGAGTTGTCTGGGTGTGGAGATCTGAGACGGTCTGCGGCCCACTTGGTTGCTTGG 556

## RESULT 8

US-10-042-894A-15  
; Sequence 15, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 899  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (89)....(424)  
US-10-042-894A-15

Query Match 51.0%; Score 686; DB 14; Length 899;  
Best Local Similarity 88.7%; Pred. No. 4e-185;  
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;  
Qy 2 CACGAGGTGAGTCCGTACCCCTCGCGCCATAGTCCCTTCCCATACCATGTCCGACC 61  
Db 45 CACCATCGCCACCCGTACCCCTTGCTCCCATAG-----TCCCATATACCATGCGCGACC 98  
Qy 62 TCACCCCGCGGAGACCAAGTCCGCGCCACCGCGCTCCGCCAGCAAGCTGGGGCCGC 121  
Db 99 TCACCCCGCGGAGACCAAGTCCGCGCTCACCGCGCTCCGCCAGCAAGCTGGGGCCGC 158  
Qy 122 TCATCAGCGGTTCGGGCTCTTTCTAAGCCGCTCCAGSCCGCGCACCGTGGGGAGCAG 181  
Db 159 TCATCAGCGGTTCGGGCTCTTTCTAAGCCGCTCCAGSCCGCGCACCGTGGGGAGCAG 218  
Qy 182 AGGTGCGCTTCTATAGGGGTTCCTCGGCCACGCGCGCTCCCGGCCCGCATCCGAGACA 241  
Db 219 AGGTGCGCTTCTATAGGGGTTCCTCGGCCACGCGCGCTCCCGGCCCGCATCCGAGACA 278  
Qy 242 CTTCTTCCCCCGGTTTCACGGCAGCGACTCTCTCCACACCGAGGCGCAGCCCGGGGAGC 301  
Db 279 CTTCTTCCCCCGGTTTCACGGCAGCGACTCTCTCCACACCGAGGCGCAGCCCGGGGAGC 338  
Qy 302 CGCATCTCACCTCGTCTCTCGACGACTCTCTCGCGGGGTTTCAGGCGCCCTCGTGCAG 361  
Db 339 CGCATCCGTACCTCGTCTCTCGACGACTCTCTCGCGGGGTTTCAGGCGCCCTCGTGCAG 398





; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 643  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(643)  
; OTHER INFORMATION: n = A, T, C or G  
US-10-042-894A-17

Query Match 35.0%; Score 470.4; DB 14; Length 643;  
Best Local Similarity 95.9%; Pred. No. 1.1e-123;  
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;  
QY 42 TCCCATACATGTCGACCTCCACCCCGCGGAGCACCAAGTGGCGGCACCGCGCTC 101  
DB 107 TCCCATACATGTCGACCTCCACCCCGCGGAGCACCAAGTGGCGGCCTC 166  
QY 102 CGCCAGACAGTGGGCGGCTCATCGACGGCTCCGGGCTCTTCTACAGCGCGTCCAGGC 161  
DB 167 CGCCAGACAGTGGGCGGCTCATCGACGGCTCCGGGCTCTTCTACAGCGCGTCCAGGC 226  
QY 162 CGCGGACCGTGGGAGCACAGAGTGCCTTCTATAGAGGCTTCTCGCGCCACCGCGCGT 221  
DB 227 CGCGGACCGTGGGAGCACAGAGTGCCTTCTATAGAGGCTTCTCGCGCCACCGCGCGT 286  
QY 222 CGCGGCGCGATCCGAGACACCTTCTTCCCGGTTTCCAGGACCGGACCTCTCCCGAC 281  
DB 287 CGCGGCGCGATCCGAGACACCTTCTTCCCGGTTTCCAGGACCGGACCTCTCCCGAC 346  
QY 282 CGAGGCGGACCGCGGAGCGGACCTCTACCTGCTCTCGAGGACCTCTCGCGGGTT 341  
DB 347 CGAGGCGGACCGCGGAGCGGACCTCTCGGACCTCTCGGACGACCTCTCGCGGGTT 406  
QY 342 TCAGGCGGCTCGTCCGAGACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCGCC 401  
DB 407 TGAGGCGGCTCGTCCGAGACATCAAGATCGCGCCATCACTGCGCCACCGAGTTCGCC 466  
QY 402 GGAGCCCTACATCGCAAGTGTCTGCCAAGACCGCGGACCAAGAGCGTTCTGCTCGG 461  
DB 467 GGAGCCCTACATCGNCAAGTACTGCGCCAAAGACCGCGGACCAAGAGCGTTCTGCTCGG 526  
QY 462 ATTCGCGTCTCGGCGTCCGAGTGTCTCGCGCCCGAGGCGCGTGTGCGGACCGAGCG 521  
DB 527 ATTCGCGTCT--TGCGTCCGAGTGTCTCGGCGCCCGAGGCGCGTGTGCGGACCGAGCG 584  
QY 522 -CCGAGAGTGAAGGCATGGACA--CGCGCGGCTCCGCGCGTGTCTCGG 570  
DB 585 CCGCGGGGTGAAGCTATGGACCCCGTCCGNGNCCGCGGNGTGTCTCGG 635

RESULT 11  
US-10-437-963-4649  
; Sequence 4649, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 4649  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_11513C.1  
US-10-437-963-4649

Query Match 22.9%; Score 308; DB 19; Length 696;  
Best Local Similarity 69.6%; Pred. No. 2.4e-77;  
Matches 451; Conservative 0; Mismatches 185; Indels 12; Gaps 2;  
QY 286 GCGCAGCCCGGGAGCGGCATCTCCTCCTCGCTCTCGACGACCTCTCTCGCGGGTTTCAG 345  
DB 49 GCGTCCGCGACAGCTGGGCGGCTCTGTCGACGCGGAGGGCTCTTCTACAGGCCCTC 108  
QY 346 GCGCCTCTGTCGACAGATCAAGATCGCGCCATCACTGTCGCCACCGAGTTTCGCGGAG 405  
DB 109 CAGGCGGGAGCGCGGGAGCACAGGCGGCTTTTACGCGCGTTTACCGCGCACCGC 168  
QY 406 CCCTACATCGCCAGTGTCTGCGCAAGACCGCGGAGCACGAGCGTTTCTGCTCGATTTC 465  
DB 169 GCGGTCGCGCGCGGCTGCTCGCCAAAGACCGCGGAGCACAGCGCGCTCTCTCGGCTTC 228  
QY 466 GCGGTCCTCGGCGTCCGAGTGTCTC---GGCCCGAGGCGCGGCTGTGCGGACGAGCGC 522  
DB 229 CGGTCCTCGGCGTCCGCGTGTGTCGATCCCGGGGCGCGGCTGTGCGGCGCGGACCGG 288  
QY 523 CGGAGGTGAAGGCGCATCGACACCGCCCGGCTGTCGCCGCTGCTCCGCGCTACGTGTC 582  
DB 289 TCGAGCTGAAGGGGATCGACCGCGCGGGTTCGCGCGGCTGCTCCGCGCTACGTGTC 348  
QY 583 TCCGTTGCGGAGGAGGATGGAATGCTGCTGCGCTCGCGCGGCTGTACGGAGGAAAGGT 642  
DB 349 ACGGCGCGCGGCGCGGCTGGAATGCTGCGCTGCGCGCGGCTGTACGGAGGAGGAGC 408  
QY 643 GGAGTCTTGTTCACAGCTCGCGGAGCTCAAGGCGTGTTCGAGGAGCAGACTCTGTTCCAC 702  
DB 409 GGCGTCTGCTGCTGCTGCTGCGGAGCTCAAGGCGTGTTCGAGGAGCAACCTGTACAC 468  
QY 703 TTCTACTCGGCGTCAATTTCTTGGGCT-----ATGATGCTGCTGCTGCTGCTGCTGCT 753  
DB 469 TTCTACTCGGCGTCAATTTCTGTTGCGCTACGACGCAATGCGGCGCGCGCTGCTGCC 528  
QY 754 GCGGAGATGGGGTGGGTCGCGTGAAGTGTGAGTGTGAGTGTGCGCATGTGCGCGAGGT 813  
DB 529 GGAGTGAAGCGCGGCTGTAAAGGTTGAAGTGTGAGTGTGCGCATGTGCGCATGGG 588  
QY 814 GATGGGCTGATTGACCAACAACTTCTCGGCGGCGCTCTGCTGCTGCTGATCAAGTTCGTTCT 873  
DB 589 GACGGGCTGATTGACCAACAACTTCTTGGCGGCGCTCTGCTGCTGCTCATAGTTTCATCGC 648  
QY 874 GACATTGTTCCGAGACTCTCTATACGAGCGCTTTGGGCTCTTTCTTAA 921  
DB 649 GACATTGTTCCGAGGTTTACCGAGAAGGCGCTTTCAGATCATTTCTTGA 696

RESULT 12  
US-10-042-894A-18  
; Sequence 18, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Caboon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase  
; TITLE OF INVENTION: Genes and Uses Thereof



```

; LOCATION: (17):(255)
; OTHER INFORMATION: unsure at all n locations

```

QY 845 GGCTCTGCTCGCTGATCAAGTTC--GTTTCTGACATTGTTCCGGAGACTCCTCATACG-C 901  
Db 181 AGCTCTGCTAGCTGATCAAGTTCCTGTTCTTGACATTGTTCCAGANACTCCTTAGACGCC 240  
QY 902 AGCCTTTGGGTCTCTTTAAAGAGA 925  
Db 241 AGCCTTTGGGTCTCTTTAAAAA 264

RESULT 15  
US-10-042-894A-11  
; Sequence 11, Application US/10042894A  
; Publication No. US20030009011A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Jinrui  
; APPLICANT: Beach, Larry  
; APPLICANT: Wang, Hongyu  
; APPLICANT: Rafalski, Antoni J.  
; APPLICANT: Cahoon, Rebecca E.  
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase  
; FILE OF INVENTION: Genes and Uses Thereof  
; FILE REFERENCE: 1286  
; CURRENT APPLICATION NUMBER: US/10/042,894A  
; CURRENT FILING DATE: 2002-01-09  
; PRIOR APPLICATION NUMBER: US 60/261,465  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 1195  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (116)...(1048)  
US-10-042-894A-11

Query Match 13.7%; Score 184.6; DB 14; Length 1195;  
Best Local Similarity 55.6%; Pred. No. 4.2e-42;  
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY 56 CGGACCTCCACCGCGGAGACCAAGTCGCGGCCACCGCCTCCGCCAGCAAGCTGG 115  
Db 114 CATGCTCAAGTCCCGATCATCAAGTCGCGGTCAACCGGAGACGGGGGAAAGCTGG 173  
QY 116 GCCCGCTCATCGACGCTCGGCTCTTCTACAAGCGCTCCAGGCGCGGACCGTGGGG 175  
Db 174 GGCCACTGGTGATGATTGGGCGGCTTCTATAAGCCTCTCCAGAGCGATCATCGGGAG 233  
QY 176 AGCAGAGTGGCTTCTATAGGGCTTCTCCGCCACCGCGCGTCCCGGCCCGATCC 235  
Db 234 ACACGGAAGTGGCTTTACGAGTCAATCTATTCCAATACCGAGATCCCGGTCACATTC 293  
QY 236 GAGACACCTTCTCCCGGTTCCAGGCAACCGACTCTCTCCCGAGGCGCAGCCCG 295  
Db 294 G---CAAAATCTTCTCTGCTTTCAGGAACTAAGACTAT---TGAGGCGTCTGATG 347  
QY 296 GGGAGCGCGATCTCACCTCGTCTCGACGACTCTCTCGCGGGTTTCAGGCGCCTGCG 355  
Db 348 CGGGTCTCAACCTCACTGGTTCTGGAGATCTCGTCTCGGTCGCACGAACCCATCTC 407  
QY 356 TCGCAGACATCAAGATCGCGGCATCACTGGCCACCGAGTTTCGCGGAGCCCTACATCG 415  
Db 408 TCATGACATCAAGACTGGATCCAGAACATGGTATCCGGAGGCTCTGAGGAGTACATCC 467  
QY 416 CCAAGTGCCTCCCGAGGACCGGGACCGAGCGTCTGCTCGGATTCGCGTCTCGG 475  
Db 468 AAAAGTGCCTTAGAGAAAGATCGAAATAGCACAAAGCGTTTCATTGGGTTTATAG 527  
QY 476 GCGTCCGAGTCTCGG--CCCGAGGGCGCGTGTGGCGGACGGAGCCCGGAGGTGA 532  
Db 528 GGCTAAGGGTATCAAAATAGGAAGCTGGATTTTGGCAACCTGAGAAGAGTTGTTT 587

QY 533 AGCCCATGGACACCCCGCGGTCCCGCGGTCTCCGGCGCTACGTGTCTATCC----- 585  
Db 588 ATAGCTTTAATGCGGACCGGTGTCAAGTCCGGCTCTGAGGAAGTTTGTCTTCCAACTTGT 647  
QY 586 -----GTTGCCGACGAGGGGATGGACTGTGCGCTCGCCGCGCGGTGTACGGAGGAAAAAG 640  
Db 648 CTCCTGGGTCCAAATGTGGATCGGATTTGTTGATATGCATCAAAAGTTTACTGTACCCGG 707  
QY 641 GTGGAGTCTTGTACAGCTGCGCGAGCTCAAGGCGTGGTTCGAGGAGCAGACTCTGTTC 700  
Db 708 GTGGAAATTTTGGCACAATTGCTTCAAGCTGAAGGAATGGTTTGAGGTTCAGACGAATATC 767  
QY 701 ACTTCTACTCGGCTCGATTCTTCTGGGCTATGATGCTGCTCAGTTCGAGCAGGCGGAG 760  
Db 768 ACTTCTATTTCTTGTCTCACTCATTTATATATGACAGGGAGTCTGCTTTGGACGGC--- 823  
QY 761 ATGGGGTGGGTGACGCTGAAGCTGGTGGACTTTGCCCATCTGGCCGAGGGTGTATGGG 820  
Db 824 --TGTGCACACCCGAAAGTTAACTGGTGGACTTTGCAATGTGATGGATGGCCACGGCG 881  
QY 821 TGATTGACCAACAATTCCTGGCGGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACATTG 880  
Db 882 TGATCGATCACAACTTCTTGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941  
QY 881 TTCCGGA 887  
Db 942 CTGATGA 948

Search completed: June 17, 2005, 04:00:54  
Job time : 893 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 22:26:07 ; Search time 4769 Seconds  
(without alignments)  
10727.276 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hic:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gsei:.\*  
9: gb\_gsei2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description        |
|------------|-------|-------------|--------|------------|--------------------|
| 1          | 921   | 68.5        | 960    | CC344833   | CG367966 OG0FJ44TH |
| 2          | 839.4 | 62.5        | 923    | CG240397   | CG367979 OG0FJ44TV |
| 3          | 762.2 | 56.7        | 818    | CG453578   | CG367979 OG0FJ44TV |
| 4          | 752   | 56.0        | 873    | CC344824   | CG284580 OG0FJ44TV |
| 5          | 644.8 | 48.0        | 867    | CG724960   | CG073494 PUJCG54TB |
| 6          | 620   | 46.1        | 801    | CA134480   | CG073494 PUJCG54TB |
| 7          | 613.2 | 45.6        | 719    | CG284572   | CG073494 PUJCG54TB |
| 8          | 589.4 | 43.9        | 591    | CG776236   | CG073494 PUJCG54TB |
| 9          | 564.6 | 42.0        | 656    | CA245555   | CG073494 PUJCG54TB |
| 10         | 559   | 41.6        | 722    | CA130685   | CG073494 PUJCG54TB |
| 11         | 557.8 | 41.5        | 3374   | AY109355   | CG073494 PUJCG54TB |
| 12         | 555   | 41.3        | 671    | CA202412   | CG073494 PUJCG54TB |
| 13         | 540.8 | 40.2        | 722    | CG724950   | CG073494 PUJCG54TB |
| 14         | 535.6 | 39.9        | 593    | BZ774817   | CG073494 PUJCG54TB |
| 15         | 525.6 | 39.1        | 744    | CA130686   | CG073494 PUJCG54TB |
| 16         | 510.8 | 38.0        | 960    | CG073495   | CG073494 PUJCG54TB |
| 17         | 500   | 37.2        | 907    | CG240387   | CG073494 PUJCG54TB |
| 18         | 485   | 36.1        | 955    | CG384766   | CG073494 PUJCG54TB |
| 19         | 440.4 | 32.8        | 1131   | CK211413   | CG073494 PUJCG54TB |
| 20         | 435.6 | 32.4        | 794    | 2 BF617713 | CG073494 PUJCG54TB |
| 21         | 426.2 | 31.7        | 575    | CG734025   | CG073494 PUJCG54TB |
| 22         | 423   | 31.5        | 551    | BH408067   | CG073494 PUJCG54TB |
| 23         | 420.4 | 31.3        | 619    | CA204213   | CG073494 PUJCG54TB |
| 24         | 376.2 | 28.0        | 842    | BZ736665   | CG073494 PUJCG54TB |

|      |       |       |      |     |          |           |           |
|------|-------|-------|------|-----|----------|-----------|-----------|
| 25   | 376.2 | 28.0  | 866  | 9   | CG367966 | CG367966  | QG0FJ44TH |
| C 26 | 376.2 | 28.0  | 939  | 9   | CG367979 | QG0FJ44TV |           |
|      | 27    | 372.8 | 27.7 | 588 | 5        | BU999708  | HL15H04r  |
| C 28 | 351   | 26.1  | 927  | 9   | CG284580 | OGWIF61TV |           |
|      | 29    | 339.2 | 25.2 | 876 | 9        | CG073494  | PWJCG54TB |
| 30   | 334.4 | 24.9  | 431  | 9   | CG776505 | 1123005E0 |           |
| 31   | 334.2 | 24.9  | 557  | 2   | BF277833 | GA_EB003  |           |
| 32   | 333.6 | 24.8  | 448  | 8   | CG028437 | 3591_1_10 |           |
| C 33 | 332.8 | 24.8  | 1099 | 7   | CK215908 | FGAS02788 |           |
| C 34 | 324.8 | 24.2  | 1123 | 7   | CK214494 | FGAS02642 |           |
| 35   | 319   | 23.7  | 453  | 4   | BI956405 | HVSMEN000 |           |
| 36   | 311.2 | 23.2  | 433  | 4   | BM099879 | EBe001_SQ |           |
| C 37 | 310.4 | 23.1  | 317  | 9   | CG194403 | PWJG38TB  |           |
| 38   | 309.8 | 23.1  | 598  | 9   | CG286200 | OGWAV62TV |           |
| 39   | 298   | 22.2  | 875  | 9   | CG673013 | QGUFJ19TV |           |
| 40   | 297.2 | 22.1  | 620  | 5   | BQ620191 | TaLr1176F |           |
| 41   | 293.6 | 21.8  | 409  | 4   | BM099878 | EBe001_SQ |           |
| 42   | 288.8 | 21.5  | 478  | 4   | BG417083 | HVSMEX001 |           |
| 43   | 277   | 20.6  | 306  | 8   | CG023125 | 3591_1_33 |           |
| C 44 | 267.2 | 19.9  | 606  | 5   | BU051170 | 1111039E0 |           |
| 45   | 250.4 | 18.6  | 506  | 9   | CG366904 | OGQAZ50TH |           |

## ALIGNMENTS

RESULT 1  
CC344833  
LOCUS CG344833 960 bp DNA linear GSS 16-MAY-2003  
DEFINITION OG1A020TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0359C15,  
genomic survey sequence.  
ACCESSION CC344833  
VERSION CC344833.1 GI:30814239  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
REFERENCE 1 (bases 1 to 960)  
AUTHORS Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Contact: Cathy Whitelaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1. .960  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBMA0359C15"  
/clone\_lib="ZM 0.7 1.5 KB"  
/notes="Vector: pBCSK-Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

Query Match 68.5%; Score 921; DB 8; Length 960;  
Best Local Similarity 100.0%; Pred. No. 9.1e-206;  
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 35 GTCCCTTCCCATACATGTCGAGCTCCACCGCGGAGCACCAGTCCGCGGCCACC 94  
|||||  
Db 1 GTCCCTTCCCATACATGTCGAGCTCCACCGCGGAGCACCAGTCCGCGGCCACC 60  
|||||

|    |     |   |     |
|----|-----|---|-----|
| QY | 95  | GGCCCTCCGACGAGCTGGGCCGCTCATCGAGGCTCCGGCTCTTCTTACAAGCGC    | 154 |
| Db | 61  | GGCCCTCCGCCAGCAAGCTGGGCCGCTCATCGAGGCTCCGGCTCTTCTTACAAGCGC | 120 |
| QY | 155 | TCAGGCGCGCACCGTGGGAGCAGAGTGGCTTCTATGAGGCTTCTCCGCCACG      | 214 |
| Db | 121 | TCAGGCGCGCACCGTGGGAGCAGAGTGGCTTCTATGAGGCTTCTCCGCCACG      | 180 |
| QY | 215 | CGCCGCTCCGCGCGCATCCGAGACACTTCTTCCCGGTTCCAGCGCAGCGACTCC    | 274 |
| Db | 181 | CGCCGCTCCGCGCGCATCCGAGACACTTCTTCCCGGTTCCAGCGCAGCGACTCC    | 240 |
| QY | 275 | TCCCAACAGCGCGAGCCCGGAGCGCATCTCACCCTGCTCTCGAGCACTCCCTCG    | 334 |
| Db | 241 | TCCCAACAGCGCGAGCCCGGAGCGCATCTCACCCTGCTCTCGAGCACTCCCTCG    | 300 |
| QY | 335 | CGGGGTTTCAGGCGCCCTCGTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGA | 394 |
| Db | 301 | CGGGGTTTCAGGCGCCCTCGTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGA | 360 |
| QY | 395 | GTTCCGCGAGCCCTACATCGCAAGTGTCTGCCAAGGACCGCGGACACGAGGTTTC   | 454 |
| Db | 361 | GTTCCGCGAGCCCTACATCGCAAGTGTCTGCCAAGGACCGCGGACACGAGGTTTC   | 420 |
| QY | 455 | TGCTCGGATTCCGCGTCTCCGCTCCGAGTGTCTGGCCCGAGGCGCGGTGTGCGGA   | 514 |
| Db | 421 | TGCTCGGATTCCGCGTCTCCGCTCCGAGTGTCTGGCCCGAGGCGCGGTGTGCGGA   | 480 |
| QY | 515 | CGAGCGCCGAGGTTGAAGCCATCGACACCGCGCGGCTCCGCGGCTCCGCGCT      | 574 |
| Db | 481 | CGAGCGCCGAGGTTGAAGCCATCGACACCGCGCGGCTCCGCGGCTCCGCGCT      | 540 |
| QY | 575 | AGTGTATCTCGTTCGAGAGGAGTGGATGTGCGCTCGCGCGGCGGTGTACGGAG     | 634 |
| Db | 541 | AGTGTATCTCGTTCGAGAGGAGTGGATGTGCGCTCGCGCGGCGGTGTACGGAG     | 600 |
| QY | 635 | GAAAGGTGGAGTCTTGTCACAGCTGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTC  | 694 |
| Db | 601 | GAAAGGTGGAGTCTTGTCACAGCTGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTC  | 660 |
| QY | 695 | TGTTCCACTTCTACTCGCGGTGCAATCTTCTGGGCTATGATGTGTCAGTTCGAGCAG | 754 |
| Db | 661 | TGTTCCACTTCTACTCGCGGTGCAATCTTCTGGGCTATGATGTGTCAGTTCGAGCAG | 720 |
| QY | 755 | CGGAGATGGGGTGGGTGAGCGTGAAGCTGTGTGAATTTCGCCATGTGGCGCAGGGTG | 814 |
| Db | 721 | CGGAGATGGGGTGGGTGAGCGTGAAGCTGTGTGAATTTCGCCATGTGGCGCAGGGTG | 780 |
| QY | 815 | ATGGGTGATTGACCAAACTTCTGGGCGGCTCTGCTGATCAAGTTCTGTTCTG      | 874 |
| Db | 781 | ATGGGTGATTGACCAAACTTCTGGGCGGCTCTGCTGATCAAGTTCTGTTCTG      | 840 |
| QY | 875 | ACATTGTTCCGGAGACTCTCTATACGAGCTTTGGGTCTCTTTAAGAGAGGATCTCG  | 934 |
| Db | 841 | ACATTGTTCCGGAGACTCTCTATACGAGCTTTGGGTCTCTTTAAGAGAGGATCTCG  | 900 |
| QY | 935 | CATTTCGATTGATAACAAAG                                      | 955 |
| Db | 901 | CATTTCGATTGATAACAAAG                                      | 921 |

RESULT 2  
CG240397  
LOCUS  
DEFINITION  
CG240397 923 bp DNA linear GSS 22-AUG-2003  
genomic survey sequence.  
ACCESSION  
CG240397  
VERSION  
CG240397.1  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

|                             |   |
|-----------------------------|---|
| REFERENCE<br>AUTHORS        | clade; Panicoideae; Andropogoneae; Zea.<br>1 (bases 1 to 923)<br>Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,<br>Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,<br>Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.<br>Consortium for Maize Genomics<br>Unpublished (2002)<br>Other GSSs: OGYCI27TH<br>Contact: Cathy Whitelaw<br>TIGR |
| TITLE<br>JOURNAL<br>COMMENT | 9712 Medical Center Drive, Rockville, MD 20850, USA<br>Tel: 301-838-5843<br>Fax: 301-838-0208<br>Email: whitelaw@tigr.org<br>Seq primer: TP<br>Class: sheared ends.   |
| FEATURES<br>source          | Location/Qualifiers<br>1..923<br>/organism="Zea mays"<br>/mol_type="genomic DNA"<br>/strain="B73"<br>/db_xref="taxon:4577"<br>/clone="ZMMBMA078BE05"<br>/clone_lib="ZM 0.7 1.5_KB"<br>/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb<br>methylation filtered genomic DNA library"   |
| ORIGIN                      |   |
| Query Match                 | 62.5%; Score 839.4; DB 9; Length 923;   |
| Best Local Similarity       | 97.3%; Pred. No. 1.4e-186;  |
| Matches                     | 875; Conservative 0; Mismatches 21; Indels 3; Gaps 2;   |
| QY                          | 42 TCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCGCACCGGCGCTC 101  |
| Db                          | 24 TCCCATACCATGTCCGACCTCCACCGCGGAGCACCAAGTCGCGGCTC 83   |
| QY                          | 102 CGCCAGCAAGCTGGGCGCGCTCATCGAGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 161   |
| Db                          | 84 CGCCAGCAAGCTGGGCGCGCTCATCGAGGCTCCGGCTCTTCTACAAGCGCTCCAGGC 143  |
| QY                          | 162 CGCGCACCTGTGGGAGCAGAGGTCCCTTCTATGAGGCTTCTCCGCCACCGCGCGCT 221  |
| Db                          | 144 CGCGCACCTGTGGGAGCAGAGGTCCCTTCTATGAGGCTTCTCCGCCACCGCGCGCT 203  |
| QY                          | 222 CCGGCGCGCATCCGAGACACCTTCTCCCGGTTCCAGCGACGCGACTCTCTCCAC 281  |
| Db                          | 204 CCGGCGCGCATCCGAGACACCTTCTCCCGGTTCCAGCGACGCGACTCTCTCCAC 263  |
| QY                          | 282 CGAGCGCAGCCGCGGAGCGCATCTCACCTCGTCTTCCAGCAGCTCTCGCGGGTT 341  |
| Db                          | 264 CGAGCGCAGCCGCGGAGCGCATCTCGCACCTCTCTCGAGCAGCTCTCGCGGGTT 323  |
| QY                          | 342 TCAGCGCGCTCGTTCGAGACATCAAGATCGCGCGCATCAGTGGCCACCGAGTTCGCC 401   |
| Db                          | 324 TCAGCGCGCTCGTTCGAGACATCAAGATCGCGCGCATCAGTGGCCACCGAGTTCGCC 383   |
| QY                          | 402 GGAGCGCTTACATCGCCAGTGCCTCGCAAGACCGCGGACACGAGGTTCTGCTCG 461  |
| Db                          | 384 GGAGCGCTTACATCGCCAAAGTACCTCGCCAAAGACCGCGGACACGAGGTTCTGCTCG 443  |
| QY                          | 462 ATTCCGCTCTCCGCGTCCGAGTGTCTCGGCGCCGAGGCGCGCTGTGCGGACGAGCG 521  |
| Db                          | 444 ATTCCGCTCT--TGGTCCGAGTGTCTCGGCGCCGAGGCGCGCTGTGCGGACGAGCG 501  |
| QY                          | 522 CCGGAGGTGAAGGCATGGACACCGCGCGCTTCGCGCGGCTCTCCGCGGCTTACGTGTC 581  |
| Db                          | 502 CCGGAGGTGAAGGCTATGGACACCGTTCGCGCTCCGCGCGCTCTCCGCGGCTTACGTGTC 561  |
| QY                          | 582 ATCCG--TTGCCGACGAGGGATGGATGTGCGCTCGCCGCGGCTGTACGAGGAAAG 640   |
| Db                          | 562 ATCCGCTTTCGCCACGAGGGGATGGATGTGCGCTCGCGCGGCGGTGTACGAGGAAAG 621   |
| QY                          | 641 GTGGAGTCTTGTCAACAGCTCGCGAGCTCAAGGCGTGGTTTCGAGGAGCAGACTCTCTTCC 700   |

|                       |            |   |   |
|-----------------------|------------|---|---|
| Db                    | 622        | GTGAGTCTTTGTCAAGCTGCGCAGCTCAAGGCATGTTTGGAGGAGCAGACTCTGTTC     | 681   |
| Qy                    | 701        | ACTTCTACTCGGCGTCGATTCCTCTGGGCTATGATGCTGTCAGTCCGACAGCGCGAG     | 760   |
| Db                    | 682        | ACTTCTACTCGGCGTCGATTCCTCTGGGCTATGATGCTGTCAGTCCGACAGCGCGAG     | 741   |
| Qy                    | 761        | ATGGGGGTGGGGTGACCGTGAAGCTGGTGGACTTTTGCCCATGTGCGCCGAGGGTGATGGG | 820   |
| Db                    | 742        | GTGGGGGTGGGGTAACAGTGAAGCTGGTGGACTTTTGCCCATGTGCGCCGAGGGTGATGGG | 801   |
| Qy                    | 821        | TGATTGACCACAACCTTCCTCGGCGGGCTCTGCTCGCTGATCAAGTTCGTTCTGACATTG  | 880   |
| Db                    | 802        | TGATTGACCACAACCTTCCTCGGCGGGCTCTGCTCGCTGATCAAGTTCGTTCTGACATTG  | 861   |
| Qy                    | 881        | TTCCGGAGACTCCCTCATACGCGAGCCTTTGGGTCTCTTTAAGAGAGAGATCCTGGCATTT | 939   |
| Db                    | 862        | TTCCAGAGACTCCTTAGACGCGAGCCTTTGGGTCTCTTTAAGAGAGAGATCCTGGCATTT  | 920   |
| RESULT 3              |            |   |   |
| CG453578              |            |   |   |
| LOCUS                 | CG9AD40TV  | ZM_0.7_1.5_KB   | Zea mays genomic clone ZMMBMa0849H08, linear GSS 17-SEP-2003  |
| DEFINITION            |            |   | genomic survey sequence.  |
| ACCESSION             | CG453578   |   |   |
| VERSION               | CG453578.1 | GI:34838578   |   |
| KEYWORDS              | GSS.       |   |   |
| SOURCE                | Zea mays   |   |   |
| ORGANISM              | Zea mays   |   |   |
| REFERENCE             |            |   |   |
| AUTHORS               |            |   | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  |
| TITLE                 |            |   | 1. (bases 1 to 818)   |
| JOURNAL               |            |   | Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Reenick,A., Frazer,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.  |
| COMMENT               |            |   | Consortium for Maize Genomics<br>Unpublished (2002)<br>Other_GSSs: OG9AD40TH<br>Contact: Cathy Whitelaw<br>TIGR   |
| FEATURES              |            |   |   |
| source                |            |   | 1. 818<br>/organism="Zea mays"<br>/mol_type="genomic DNA"<br>/strain="B73"<br>/db_xref="taxon:4577"<br>/clone="ZMMBMa0849H08"<br>/clone_lib="ZM 0.7 1.5 KB"<br>/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library" |
| ORIGIN                |            |   |   |
| Query Match           |            | 56.7%   | Score 762.2; DB 9; Length 818;  |
| Best Local Similarity |            | 99.6%   | Pred. No. 2e-168;   |
| Matches               | 764;       | Conservative 0;   | Mismatches 3; Indels 0; Gaps 0;   |
| Qy                    | 8          | GTCAAGTCGTCACCCCTCGGGCCCATAGTCCCGCTTCCCATACCATGTCCGACCTCCACC  | 67  |
| Db                    | 52         | GTCAAGTCGTCACCCCTCGGGCCCATAGTCCCGCTTCCCATACCATGTCCGACCTCCACC  | 111   |
| Qy                    | 68         | CGCCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCGCAGCAAGTCGCGCCGCTCATCG    | 127   |
| Db                    | 112        | CGCCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCGCAGCAAGTCGCGCCGCTCATCG    | 171   |
| Qy                    | 128        | ACGGTCCGGCGCTCTTCTACAGCGCGCTCCAGGGCGGGACCGTGGGGAGCACGAGGTGC   | 187   |

|            |   |  |  |     |
|------------|---|--|--|-----|
| Db         | 172   |  | ACGGCTCCGGCCTCTTCAAGCCGCTCAGGCGCGGACCGTGGGAGCACAGGTCG          | 231 |
| Qy         | 188   |  | CTTCTATATAGGCGTTCTCCGCCACACGCCCGTCCCGGCCGCGCATCCGAGACACCTTCT   | 247 |
| Db         | 232   |  | CTTCTATATAGGCGTTCTCCGCCACACGCCCGCGTCCCGGCCGCGCATCCGAGACACCTTCT | 291 |
| Qy         | 248   |  | TCGCCCGGTTCCAGGCAACGCGACTCTCTCCCAACGAGGGCGACGCCGGGAGCGGCATC    | 307 |
| Db         | 292   |  | TCGCCCGGTTCCAGGCAACGCGACTCTCTCCCAACGAGGGCGACGCCGGGAGCGGCATC    | 351 |
| Qy         | 308   |  | CTACCTCTGTCGAGACCTCTCTCCGGGGTTTCAAGGCGCCTCGTTCGACAGACATCA      | 367 |
| Db         | 352   |  | CTACCTCTGTCGAGACCTCTCTCCGGGGTTTCAAGGCGCCTCGTTCGACAGACATCA      | 411 |
| Qy         | 368   |  | AGATCGGCGGCATCACCTGGGCCACCGAGTTTCGCCGAGCCCTACATCGCCAAAGTGCCTCG | 427 |
| Db         | 412   |  | AGATCGGCGGCATCACCTGGGCCACCGAGTTTCGCCGAGCCCTACATCGCCAAAGTGCCTCG | 471 |
| Qy         | 428   |  | CCAAGACCCGCGGACCAACGAGGTTCTGTTCGATTTCGGGTTCTCGGCGTTCGAGTCG     | 487 |
| Db         | 472   |  | CCAAGACCCGCGGACCAACGAGGTTCTGTTCGATTTCGGGTTCTCGGCGTTCGAGTCG     | 531 |
| Qy         | 488   |  | TCGGCCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGGAGGTTCGAGCCATGGACACCG     | 547 |
| Db         | 532   |  | TCGGCCCCGAGGGCGCGTGTGGCGGACGAGCGCCCGGAGGTTCGAGCCATGGACACCG     | 591 |
| Qy         | 548   |  | CCGCGCTCCGCGGCTCTCCGGCGTACGTTGTCATCCGTTGCCGACGAGGGGATGGACT     | 607 |
| Db         | 592   |  | CCGCGCTCCGCGGCTCTCCGGCGTACGTTGTCATCCGTTGCCGACGAGGGGATGGACT     | 651 |
| Qy         | 608   |  | GTGCGCTCGCGCGCGCGTGTACGAGGAAAAGGTGGAGTCTTGTCAAGTTCGCGGAGC      | 667 |
| Db         | 652   |  | GTGCGCTCGCGCGCGCGTGTACGAGGAAAAGGTGGAGTCTTGTCAAGTTCGCGGAGC      | 711 |
| Qy         | 668   |  | TCAAGGCGTGTTCGAGGACGAGTCTGTTCGACTTCTACTCGCGCTCGATTCTTCTCG      | 727 |
| Db         | 712   |  | TCAAGGCGTGTTCGAGGACGAGTCTGTTCGACTTCTACTCGCGCTCGATTCTTCTCG      | 771 |
| Qy         | 728   |  | GCTATGATGCTGCTGCAGTTCGACGACGCGGAGATGGGGTGGGGTG                 | 774 |
| Db         | 772   |  | GCTATGATGCTGCTGCAGTTCGACGACGCGGAGATGGGGTGGGGTG                 | 818 |
| RESULT 4   |   |  |  |     |
| CC344824/C |   |  |  |     |
| LOCUS      |   |  |  |     |
| DEFINITION | CGIAQ20TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMA0359C15,   |  |  |     |
| ACCESSION  | CG344824  |  |  |     |
| KEYWORDS   | GSS.  |  |  |     |
| SOURCE     | Zea_mays  |  |  |     |
| ORGANISM   | Zea_mays  |  |  |     |
| REFERENCE  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  |  |  |     |
| AUTHORS    | Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Frazer,C.M., Buddman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.   |  |  |     |
| TITLE      | Consortium for Maize Genomics   |  |  |     |
| JOURNAL    | Unpublished (2002)  |  |  |     |
| COMMENT    | Contact: Cathy Whitelaw<br>TCGR<br>9712 Medical Center Drive, Rockville, MD 20850, USA<br>Tel: 301-838-5843<br>Fax: 301-838-0208<br>Email: whitelaw@tigr.org<br>Seq primer: TR<br>Class: sheared ends.<br>Location/Qualifiers<br>1..873 |  |  |     |
| FEATURES   | source  |  |  |     |

```
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM0.7.1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      56.0%; Score 752; DB 8; Length 873;
Best Local Similarity 100.0%; Pred. No. 5.1e-166;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 CTCGCGCCACGCGCGTCCGCGCCCGCATTCGAGACACTTCTTCCCGCGGTTCACGG 263
DB 873 CTCGCGCCACGCGCGTCCGCGCCCGCATTCGAGACACTTCTTCCCGCGGTTCACGG 814
QY 264 CACGCGACTCTCTCCACACGAGCGCAGCCCGGGAGCGCATCTCTCACTCGTCTCGA 323
DB 813 CACGCGACTCTCTCCACACGAGCGCAGCCCGGGAGCGCATCTCTCACTCGTCTCGA 754
QY 324 CGACTCTCTCGGGGGTTTCAGCGCCCTGCGTCGCAGACATCAAGATCGGGCCATCAC 383
DB 753 CGACTCTCTCGGGGGTTTCAGCGCCCTGCGTCGCAGACATCAAGATCGGGCCATCAC 694
QY 384 GTGGCCACCGAGTTCCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCAAAGAGCGGGGAC 443
DB 693 GTGGCCACCGAGTTCCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCAAAGAGCGGGGAC 634
QY 444 CACGAGCGTTCTGCTCGGATTCGCGTCTCCGCGTCCGAGTCTGCGGCCCGAGGGGCG 503
DB 633 CACGAGCGTTCTGCTCGGATTCGCGTCTCCGCGTCCGAGTCTGCGGCCCGAGGGGCG 574
QY 504 CTTGTGGCGGACGAGCGCCGAGAGTGAAGGCCATGAGACACCGCGCGTCCGCGCGGT 563
DB 573 CTTGTGGCGGACGAGCGCCGAGAGTGAAGGCCATGAGACACCGCGCGTCCGCGCGGT 514
QY 564 GCTCGGCGCTAGTGTGATCGTTGCCGACGAGGGGATGAGTGTGCGCTCGCGCGGC 623
DB 513 GCTCGGCGCTAGTGTGATCGTTGCCGACGAGGGGATGAGTGTGCGCTCGCGCGGC 454
QY 624 GTGTACGAGGAAAGTGGAGTCTGTGCACAGCTCGCGGAGCTCAAGCGTGGTTGA 683
DB 453 GTGTACGAGGAAAGTGGAGTCTGTGCACAGCTCGCGGAGCTCAAGCGTGGTTGA 394
QY 684 GGAGCAGACTCTGTTCCACTTCTACTCGGCGTCGATTTCTTCTGGGCTATGATCTGCTGC 743
DB 393 GGAGCAGACTCTGTTCCACTTCTACTCGGCGTCGATTTCTTCTGGGCTATGATCTGCTGC 334
QY 744 AGTCGACGACGCGGAGATGGGGGTGGGTGACCGGTGAAGCTGGTGGACTTTGCCCATGT 803
DB 333 AGTCGACGACGCGGAGATGGGGGTGGGTGACCGGTGAAGCTGGTGGACTTTGCCCATGT 274
QY 804 GSCGAGGGTGTATGGGTGATTTGACCACAACTTCTTGGGCGGCTCTGCTCGCTGATCAA 863
DB 273 GSCGAGGGTGTATGGGTGATTTGACCACAACTTCTTGGGCGGCTCTGCTCGCTGATCAA 214
QY 864 GTTCTGTTCTGACATTTGTTCCGAGACTCTCTCATAGCAGCCTTTGGGTCTCTTTAAGA 923
DB 213 GTTCTGTTCTGACATTTGTTCCGAGACTCTCTCATAGCAGCCTTTGGGTCTCTTTAAGA 154
QY 924 GAGGATCTCGGATTTTCGATTTTGATTAACAAAG 955
DB 153 GAGGATCTCGGATTTTCGATTTTGATTAACAAAG 122

RESULT 5
CC724960
LOCUS
DEFINITION
genomic survey sequence.
ACCESSION
CC724960
```

```
CC724960.1 GI:32143893
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 867)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGLAB23TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: tp
Class: sheared ends.
Location/Qualifiers
1..867
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0307C22"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      48.0%; Score 644.8; DB 9; Length 867;
Best Local Similarity 89.4%; Pred. No. 8.4e-141;
Matches 753; Conservative 0; Mismatches 27; Indels 62; Gaps 3;

QY 2 CACGAGTTCAGTCCGTCACTCCCTCGCGCCCATAGTTCCTTCCCATACCATGTCGAC 61
DB 87 CACCATCGCCACCCGTCACTCCCTTCTCTCCATAG-----TCCCATACCATGTCGAC 140
QY 62 TCCACCCCGCGGAGACCAAGTCCGCGGCCACCGCGCTCCGCCACGAGCTGGGGCCGC 121
DB 141 TCCACCCCGCGGAGACCAAGTCCGCGGCCACCGCGCTCCGCCACGAGCTGGGGCCGC 200
QY 122 TCATCGAGCGTCCGCGCTCTTCTACAAGCCGCTCCAGCGCGGACCGTGGGGAGCAG 181
DB 201 TCATCGAGCGCTCTGGCTCTTCTACAAGCCGCTCCAGCGCGGACCGTGGGGAGCAG 260
QY 182 AGTTCGCTTCTATGAGCGGTTCTCCGCCACCGCGCGTCCCGCCCGCATCCGAGACA 241
DB 261 AGTTCGCTTCTATGAGCGGTTCTCCGCCACCGCGCGTCCCGCCCGCATCCGAGACA 320
QY 242 CTTCTTCCCGCGTTCACGCGACGACTCTCTCCACCGAGCGGAGCCCGGGAGC 301
DB 321 CTTCTTCCCGCGTTCACGCGACGACTCTCTCCACCGAGCGGAGCCCGGGAGC 380
QY 302 CGCATCTCTACCTCGTCTCTCGACGACTCTCTCGGGGGTTTCAGGCGCCCTCGCTCGCAG 361
DB 381 CGCATCTCTCGTCTCTCGACGACTCTCTCGGGGGTTTCAGGCGCCCTCGCTCGCAG 440
QY 362 ACATCAAGATCGGCCCATCATCATGGCCACCGAGTTCGCCGAGCCCTACATCGCCAAAGT 421
DB 441 ACATCAAGATCGGTGCCATCAGT-----GACCCAGCGCTTCTGCTCGATTCCTCGCGTCC 464
QY 422 GCCTCGCCAAAGACCGCGGACCCAGCGGCTTCTGCTCGGATTCGCGCTCTCCGCGTCC 481
DB 465 -----GACCCAGCGCTTCTGCTCGATTCCTCGCGTCTCTCCGCGTCC 505
QY 482 GAGTCTCGGCGCCGAGGGCGCGTGTGGCGGACGAGCGCCCGGAGGTGAAGGCATGG 541
DB 506 GAGTCTCGGCGCCGAGGGCGCGTGTGGCGGACGAGCGCCCGGAGGTGAAGGCTATGG 565
```

```
QY 542 ACACCGCGCGTCCGCGCGTCTCCGCGCTACGTGTATCCG-TTGCAGAGGGG 600
Db |||
566 ACATTGTCGCGTCCGCGCGTCTCCGCGCTACGTGTATCCGCTTGCAGAGGGG 625
QY 601 ATGAGCTGTGCGTTCGCGCGCGGTGTACGAGGAAAGGTGAGTCTTTGTACAGCTG 660
Db |||
626 ATGAGCTGTGCGTTCGCGCGCGGTGTACGAGGAAAGGTGAGTCTTTGTACAGCTG 685
QY 661 CGGAGCTCAAGCGTGTGCGCGCGGTGTACGAGGAAAGGTGAGTCTTTGTACAGCTG 720
Db |||
686 CGGAGCTCAAGCGTGTGCGCGCGGTGTACGAGGAAAGGTGAGTCTTTGTACAGCTG 745
QY 721 CTTCTGGGCTATGATGCTGTCAGTGCAGTGCAGCGCGGAGATGGGGTGGGGTGAACGTG 780
Db |||
746 CTTCTGGGCTATGATGCTGTCAGTGCAGTGCAGCGCGGAGATGGGGTGGGGTGAACGTG 805
QY 781 AAGCTGTGGACTTTGCCCATGTGGCCGAGGGTGTATGGGGTGTATGACCAACTTCTTG 840
Db |||
806 AAGCTGTGGACTTTGCCCATGTGGCCGAGGGTGTATGGGGTGTATGACCAACTTCTTG 865
QY 841 GG 842
Db ||
866 GG 867

RESULT 6
CA134480
LOCUS SCJFRT1061H11.9 RT1 Saccharum officinarum cDNA clone SCJFRT1061H11
5', mRNA sequence.
CA134480
CA134480.1 GI:35021536
EST.
Saccharum officinarum
Saccharum officinarum
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 801)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 061 row: H column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .801
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFRT1061H11"
/lab_host="DH10B"
/clone_lib="RT1"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System Kit
(invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
```

construction can be obtained at  
<http://sucest.lad.ic.unicamp.br/public>

```
ORIGIN
Query Match 46.1%; Score 620; DB 6; Length 801;
Best Local Similarity 89.2%; Pred. No. 5.7e-135;
Matches 701; Conservative 0; Mismatches 82; Indels 3; Gaps 3;
QY 37 CCCTTTCCCATATGCTCCGACCTCCACCCGCGGAGCACCAAGTCCGCGCCACCGC 96
Db 16 CTTCTCGCCCTTACCATGTCGACCTCCGCGCGCGGAGCACCAAGTCCGCGCCACCGC 75
QY 97 GCCTCCGCGCAGCAAGCTGGGCGCGCTCATCGAGCTCCGCGCTCTTCTAACAAGCGCTC 156
Db 76 GCCTCCGCGCAACAAGCTGGGTCGCTCATCGAGCGCTCTGCGCTCTTCTAACAAGCGCTC 135
QY 157 CAGCGCGCGACCGTGGGAGCAGAGTCCGCTTCTATGAGGGGTCTTCGGCCACGCC 216
Db 136 CAGGTCCGGGACCGCGGGGAGCAGAGCTCGCTTCTACGAGGCGTTCTCCACCCACGCC 195
QY 217 GCGTCCCGCGCGCATCCGAGACACTTCTTCCCGCGGTTCACGAGCGTTCTCCACCCACGCC 276
Db 196 GCGTCCCGCGCGCATCCGAGACACTTCTTCCCGCGGTTCACGAGCGTTCTCCACCCACGCC 255
QY 277 CCACCGAGCGCAGCGCGGAGCGCATTCCTACCTCTCTCGACGACTCTCTCGCG 336
Db 256 CCACCGAGCGCGCGCGAGCGCATTCCTCTCTCGACGACTCTCTCGCG 315
QY 337 GGTTCACGCGCGCTCGTTCGAGACATCAAGATCGGCGCCATCAC-GTGGCCACCGAG 395
Db 316 GGGCTGGAGCGCGCTCGCTCGCGGACATCAAGATCGGCGCCATCAAGTGGCGCGCGAG 375
QY 396 TTGCG-CGGAGCGCTACATCGCCAAAGTCCCTCCGCAAGGACCGCGGACCAAGCGCTTC 454
Db 376 CTGCGGAGAGCGCTTACGTCTCAAGTCCCTGTCAAGGACCGCGGACCAAGGCAATTC 435
QY 455 TGCTCGGATTCGCGCTCTCGCGCTCCGAGTCTGCGCCCGGAGGCGCGTGTGCGGA 514
Db 436 TGCTCGGATTCGCGCTCTCGCGCTCCGCGTCTGCTCCGAGGCGCGTGTGCGGA 495
QY 515 CGAGCGCGCGAGGTGAAGCCATGAGACACCGCGCGCTCCGCGCGCTCTCGCGGCT 574
Db 496 CTGAGCGCGCGAGGTGAAGCGCTTCGACACCGCGCGCTCCGCGCGCTCTCGCGGCT 555
QY 575 ACCTGTATCCCTTCCGCGAGGAGTGGACTGTGCGCTCGCGCGCGGTGTACGAG 634
Db 556 ACCTGTATCCCTTCCGCGAGGAGTGGACTGTGCGCTCGCGCGCGGTGTACGAG 615
QY 635 GAAAGGTGAGTCTTGTTCACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGAGCAGACTC 694
Db 616 GNAAGGAGGAGTTTGTCAATTTGGCGAGCTTAAGCGTGTGTTGAGGACAGACTC 675
QY 695 TGTTCACATCTTACTCGCGCTCGATTTCTTGGGCTATGATGTCTGCTGCTGCTGCTG 754
Db 676 TGGTCCACCTTTACTCGCGGCGGATTTCTTGGGCTATGAGCTAGTGTGCTGCTGCTG 735
QY 755 GCGG-AGATGGGCGGTGAGCTGAGCTGTGAGCTTTGCCCATGTGGCGGAGGCT 813
Db 736 CCGAAGGTAAAGGTGGGTGAAGGTAAACTGGGGCCCTTTGCCCATGTGGCGGAGG 795
QY 814 GATGGG 819
Db 796 GAAGGG 801

RESULT 7
CG284572
LOCUS OSW161TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBma0584K02,
DEFINITION genomic survey sequence.
ACCESSION CG284572
VERSION CG284572.1 GI:34198786
KEYWORDS GSS.
```



|   |   |  |
|---|---|--|
| SOURCE ORGANISM   | Zea mays  |  |
| REFERENCE AUTHORS   | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.            |  |
| TITLE JOURNAL COMMENT   | 1 (bases 1 to 719)  |  |
| Other_GSSs: OGWIF61TW   | Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N. |  |
| Contact: Cathy Whitelaw   | Consortium for Maize Genomics   |  |
| Seq primer: TR  | Unpublished (2002)  |  |
| Class: sheared ends.  | Other_GSSs: OGWIF61TW   |  |
| Location/Qualifiers   | Contact: Cathy Whitelaw   |  |
| 1..719  | TIGR  |  |
| /organism="Zea mays"  | 9712 Medical Center Drive, Rockville, MD 20850, USA   |  |
| /mol_type="genomic DNA"   | Tel: 301-838-5843   |  |
| /strain="B73"   | Fax: 301-838-0208   |  |
| /db_xref="taxon:4577"   | Email: whitelaw@tigr.org  |  |
| /clone="ZMMBMA0594K02"  | Seq primer: TR  |  |
| /clone_lib="ZM_0.7_1.5_KB"  | Class: sheared ends.  |  |
| /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"  | Location/Qualifiers   |  |
| 45.6%; Score 613.2; DB 9; Length 719;   | 1..719  |  |
| Query Match 90.4%; Pred. No. 2.2e-133;  | /organism="Zea mays"  |  |
| Matches 700; Conservative 0; Mismatches 18; Indels 56; Gaps 2;  | /mol_type="genomic DNA"   |  |
| QY 50 CCATGTCGACGAGTCCACCGCGGAGCACCAGTCCGCGCCACCGCGCTCCGCCAGCA 109  | /strain="B73"   |  |
| DB 1 CCATGCCGACCTCCACCGCGGAGCACCAGTCCGCGCTCCGCCAGCA 60  | /db_xref="taxon:4577"   |  |
| QY 110 AGCTGGGCGCGCTCATCGAGGCTCCGGCTCTTTCTAAGCGCGCTCCAGGCGGCGACC 169  | /clone="ZMMBMA0594K02"  |  |
| DB 61 AGCTGGGCGCGCTCATCGAGGCTCCGGCTCTTTCTAAGCGCGCTCCAGGCGGCGACC 120   | /clone_lib="ZM_0.7_1.5_KB"  |  |
| QY 170 GTGGGAGCAGAGTCCGCTTCTATGAGCGTTCCTCGGCCACCGCGCGTCCCGGCC 229   | /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"  |  |
| DB 121 GTGGGAGCAGAGTCCGCTTCTATGAGCGTTCCTCGGCCACCGCGCGTCCCGGCC 180   | 45.6%; Score 613.2; DB 9; Length 719;   |  |
| QY 230 GCATCCGAGACACTTCTTCCCGCGGTTCCACGCGACGAGTCTCTCCCGAGGCGC 289   | Query Match 90.4%; Pred. No. 2.2e-133;  |  |
| DB 181 GCATCCGAGACACTTCTTCCCGCGGTTCCACGCGACGAGTCTCTCCCGAGGCGC 240   | Matches 700; Conservative 0; Mismatches 18; Indels 56; Gaps 2;  |  |
| QY 290 AGCCCGGGAGCGCATCT 349  | QY 50 CCATGTCGACGAGTCCACCGCGGAGCACCAGTCCGCGCCACCGCGCTCCGCCAGCA 109  |  |
| DB 241 AGCCCGGGAGCGCATCT 300  | DB 1 CCATGCCGACCTCCACCGCGGAGCACCAGTCCGCGCTCCGCCAGCA 60  |  |
| QY 350 CTTGCTCCGAGACATCAAGATCGGCGGCTCATCGTGGCCACCGAGTTCGCGGAGCCCT 409   | QY 110 AGCTGGGCGCGCTCATCGAGGCTCCGGCTCTTTCTAAGCGCGCTCCAGGCGGCGACC 169  |  |
| DB 301 CTTGCTCCGAGACATCAAGATCGGCGGCTCATCGTGGCCACCGAGTTCGCGGAGCCCT 336   | DB 61 AGCTGGGCGCGCTCATCGAGGCTCCGGCTCTTTCTAAGCGCGCTCCAGGCGGCGACC 120   |  |
| QY 410 ACATCGCCCAAGTCCCTCGCCCAAGGACCGCGGAGCACCAGCGGTTCTGCTCGGATCCCGG 469  | QY 170 GTGGGAGCAGAGTCCGCTTCTATGAGCGTTCCTCGGCCACCGCGCGTCCCGGCC 229   |  |
| DB 337 -----GACCAGAGCGTTCGCTCGGATCCCGG 365  | DB 121 GTGGGAGCAGAGTCCGCTTCTATGAGCGTTCCTCGGCCACCGCGCGTCCCGGCC 180   |  |
| QY 470 TCTCGGCGTCCGAGTCTGCGGCCCGAGGCGCGCTGTGGCGGACGAGCGCGCGGAGG 529   | QY 230 GCATCCGAGACACTTCTTCCCGCGGTTCCACGCGACGAGTCTCTCCCGAGGCGC 289   |  |
| DB 366 TCTCGGCGTCCGAGTCTGCGGCCCGAGGCGCGCTGTGGCGGACGAGCGCGCGGAGG 425   | DB 181 GCATCCGAGACACTTCTTCCCGCGGTTCCACGCGACGAGTCTCTCCCGAGGCGC 240   |  |
| QY 530 TGAAGGCATGACACCGCGCGCTCCGCGCGTCTCCGCGCGCTACGTCATCCG-TT 588   | QY 290 AGCCCGGGAGCGCATCTCTCACT 349  |  |
| DB 426 TGAAGGCATGACATTTGCGGCGTCCGCGCGTCTCCGCGCGTCTACGTCATCCGCTT 485   | DB 241 AGCCCGGGAGCGCATCT 300  |  |
| QY 589 GCCGACGAGGAGTGGACTGTGCGCTCCGCGCGCGGCTGTACGAGGAGGAAAGGTGGAGTC 648   | QY 350 CTTGCTCCGAGACATCAAGATCGGCGGCTCATCGTGGCCACCGAGTTCGCGGAGCCCT 409   |  |
| DB 486 GCCGACGAGGAGTGGACTGTGCGCTCCGCGCGCGGCTGTACGAGGAGGAAAGGTGGAGTC 545   | DB 301 CTTGCTCCGAGACATCAAGATCGGCGGCTCATCGTGGCCACCGAGTTCGCGGAGCCCT 336   |  |
| QY 649 TTGTACAGCTCGCGAGCTCAAGCGCTGTTTCGAGGAGCAGACTCTGTTCACCTTCTAC 708   | QY 410 ACATCGCCCAAGTCCCTCGCCCAAGGACCGCGGAGCACCAGCGGTTCTGCTCGGATCCCGG 469  |  |
| DB 546 TTGTACAGCTCGCGAGCTCAAGCGCTGTTTCGAGGAGCAGACTCTGTTCACCTTCTAC 605   | DB 337 -----GACCAGAGCGTTCGCTCGGATCCCGG 365  |  |
| QY 709 TCGGCGTCTGATTTCTTCTGGGCTATGCT |   |  |

```
Qy 85 GCCGGCCACCGCGCTCCGCCAGCAGCTGGGCGCGCTCATCGAGCGCTCCGGCCCTTTC 144
Db 61 GCCGGCCACCGCGCTCCGCCAGCAGCTGGGCGCGCTCATCGAGCGCTCCGGCCCTTTC 120
Qy 145 TACAAGCCGCTCCAGGCGCGCGACCGTGGGAGCAGAGGTCCGCTTCTATGAGGGCTTC 204
Db 121 TACAAGCCGCTCCAGGCGCGCGACCGTGGGAGCAGAGGTCCGCTTCTATGAGGGCTTC 180
Qy 205 TCCGCCACCGCGCGCTCCGGCGCGCGATCCGAGACACCTTCTTCCCGCGGTTCACGGC 264
Db 181 TCCGCCACCGCGCGCTCCGGCGCGCGATCCGAGACACCTTCTTCCCGCGGTTCACGGC 240
Qy 265 ACCGAGCTCTCCGCCAGGCGCGACCGCGGAGCGCGATCCTACCTCGCTCCGAC 324
Db 241 ACCGAGCTCTCCGCCAGGCGCGACCGCGGAGCGCGATCCTACCTCGCTCCGAC 300
Qy 325 GACTCTCTCCGGGGTTTCAAGGGCGCTTGGTCCGACACATCAAGATCGCGGCATCAGC 384
Db 301 GACTCTCTCCGGGGTTTCAAGGGCGCTTGGTCCGACACATCAAGATCGCGGCATCAGC 360
Qy 385 TGCCACCGAGTTCCGCGGAGCCTACATCCCAAGTCCCTGCGCAAGACCGCGGACC 444
Db 361 TGCCACCGAGTTCCGCGGAGCCTACATCCCAAGTCCCTGCGCAAGACCGCGGACC 420
Qy 445 ACAGCGTTCTGCTCGGATTCGGGTCTCCGGGTCCGAGTCCGCGCCCGAGGGGCC 504
Db 421 ACAGCGTTCTGCTCGGATTCGGGTCTCCGGGTCCGAGTCCGCGCCCGAGGGGCC 480
Qy 505 GTGTGGCGGACGAGCGCCCGAGGTGAAGCCATGACACCGCGCGCTCCCGCGTG 564
Db 481 GTGTGGCGGACGAGCGCCCGAGGTGAAGCCATGACACCGCGCGCTCCCGCGTG 540
Qy 565 CTCGGCGCTACGTTGTCATCCGTTGCGGACGAGGGGATGACTGTGCGCTC 615
Db 541 CTCGGCGCTACGTTGTCATCCGTTGCGGACGAGGGGATGACTGTGCGCTC 591
```

## RESULT 9

```
CA245555
LOCUS SCBGF5081A03.9 Saccharum officinarum FL5 Saccharum officinarum
DEFINITION cDNA clone SCBGF5081A03 5', mRNA sequence.
ACCESSION CA245555
VERSION CA245555.1 GI:35324301
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
```

```
REFERENCE 1 (bases 1 to 656)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
          Centro de Biologia Molecular e Engenhariaia Genetica
          Universidade Estadual de Campinas
          Caixa Postal 6010, 13083-970, Campinas SP, Brazil
          Tel: 55 19 3788 1137
          Fax: 55 19 3788 1089
          Email: parruda@unicamp.br
          Clone distribution: clone distribution information can be found
          through the Brazilian Clone Collection Center (BCCC) at
          http://www.bcccenter.fcav.unesp.br
          Plate: 081 row: A column: 03
          Seq primer: T7 Promoter Primer.
          Location/Qualifiers
            1..656
              /organism="Saccharum officinarum"
              /mol_type="mRNA"
              /db_xref="taxon:4547"
```

## FEATURES

## Source

```
CA130685
LOCUS SCBGF5081A03.9 R11 Saccharum officinarum cDNA clone SCCRT1004G05
DEFINITION 5', mRNA sequence.
ACCESSION CA130685
VERSION CA130685.1 GI:35013914
KEYWORDS EST.
```

```
ORIGIN
Query Match 42.0%; Score 564.6; DB 6; Length 656;
Best Local Similarity 93.7%; Pred. No. 6e-122;
Matches 599; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
Qy 37 CCCTTCCCATACCATGTCGACCTCCACCGCGCGAGCACCAAGTCGCGGCCACCGC 96
Db 18 CCTCTCGCCCTACCATGTCCGACCTCCGCGCGCGAGCACCAAGTCGCGGCCACCGC 77
Qy 97 GCCTCCCGCAGCAAGCTCGGCGCGCTCATCGAGCGCTCCGGCCCTTCTACAAAGCCGCTC 156
Db 78 GCCTCCCGCAGCAAGCTCGGCTCGCTCATCGAGCGCTTGGCCCTTCTACAAAGCCGCTC 137
Qy 157 CAGGCCGCGGACCGTGGGAGCAGCAGGTTCGCTTCTATGAGCGGTTCTCCGCCACCGC 216
Db 138 CAGGTTCGGGACCGCGCGGAGCAGCAGCTCGCTTCTACGAGCGGTTCTCACGCCACCGC 197
Qy 217 GCGTTCGGCGCGCGATTCGAGACACCTTCTTCCCGCGGTTCCACGCGACGCGACTCTC 276
Db 198 GCGTTCGGCGCGCGATTCGAGACACCTTCTTCCCGCGGTTTACGCGACGCGACTCTC 257
Qy 277 CCACCGAGCGCGCAGCCCGGAGCGCGATCCTCAGCTCGCTTCGACGACCTCTCTCGCG 336
Db 258 CCACCGAGCGCGCGCGAGCGCGATCCTCAGCTCGCTCTCGAGACCTCTCTCGCG 317
Qy 337 GGGTTTTCAGGCGCTTCGCTTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGAGT 396
Db 318 GGGCTGGAGCGCGCTTCGCTTCGCGCATCAAGATCGGCGCCATCAGTGGCGCGCGCAGC 377
Qy 397 TCGCCGAGCGCTTACATTCGCAAGTCCCTCGCCAAAGACCGCGGAGCCACGAGGTTCTG 456
Db 378 TCGCCGAGCGCTTACATTCGCAAGTCCCTCGCCAAAGTCCCTCGCGGAGCCACGAGG 437
Qy 457 CTCGGATTTCGGCTCTCCGCGCTCCGAGTCCGCGCGCGCGCGCGCGCGCGCGGACG 516
Db 438 CTCGGATTTCGGCTCTCCGCGCTCCGAGTCCGCGCTCCGCGCGCGCGCGCGCGGACT 497
Qy 517 GAGCGCCCGGAGGTGAAGGCCATCGACACCGCGCGCTCCGCGCGCTGCTCCGCGCGTAC 576
Db 498 GAGCGCCCGGAGGTGAAGGCCATCGACACCGCGCGCTCCGCGCGCTGCTCCGCGCGTAC 557
Qy 577 GTGTCTTCGCTTCGCGCGCGAGGAGTGGAGTTCGCTCCGCGCGCGCGGTTACGAGG- 635
Db 558 GTTTCATCGCTTCGCGCGCGAGGAGTGGAGTTCGCTCCGCGCGCGCGGTTACGAGG 617
Qy 636 AAAAGGTGGAGTCTTGTTCACAGCTCGCGCGAGCTCAAGGC 674
Db 618 AAAAGGTGGAGTCTTGTTCACAGCTCGCGCGAGCTTAAAGC 656
```

## RESULT 10

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

```

SOURCE      Saccharum officinarum
ORGANISM    Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE   1 (bases 1 to 722)
AUTHORS     Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE       The libraries that made SUCEST
JOURNAL     Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Setadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bccccenter.fcav.unesp.br
            Plate: 004 row: G column: 05
            Seq primer: T7 Promoter Primer.
            Location/Qualifiers
                1..722
                /organism="Saccharum officinarum"
                /mol_type="mRNA"
                /db_xref="taxon:4547"
                /clone="SCCCTT1004G05"
                /lab_host="DH10B"
                /clone_lib="RT1"
                /note="Organ: Root tips (0.3cm-long) from adult plants;
            Vector: pSPori1; Site_1: SalI; Site_2: NotI; An
            unidirectional cDNA library generated from [Root tips
            (0.3cm-long) from adult plants]. cDNA was prepared from
            polyA+ mRNA using SuperScript plasmid System Kit
            (Invitrogen). The double-strand cDNAs were fractionated
            in a sepharose CL-2B 40cm-columns and fragments sizing
            between 0.8 and 1.5 Kb were directionally cloned into the
            vector. Details of each source of RNA and library
            construction can be obtained at
            http://sucest.lad.ic.unicamp.br/public"

FEATURES             source
     source
     1..3374
     /organism="Zea mays"
     /mol_type="mRNA"
     /db_xref="MaizeDB:632179"
     /db_xref="taxon:4577"
     /clone_lib="Maize Mapping Project/DuPont Cornsensus
     Library"
     /note="this sequence is part of a project of EST
     assemblies resulting from the application of public
     contigs to seed Dupont contigs; this resource was
     assembled by Dupont as part of a collaboration for the
     overgo addressing of BACs in conjunction with the Maize
     Mapping Project"

ORIGIN
Query Match      41.6%; Score 559; DB 6; Length 722;
Best Local Similarity 90.3%; Pred. No. 1.3e-120;
Matches 639; Conservative 0; Mismatches 60; Indels 4; Gaps 4;
37  CCCTTTCCCATACCATGTCGACCTCCACCCGCGGAGCACCAGTCCGCGCCACCGC 96
   |||
   |||
   |||
16  CCTCTCGCCTCACCATGTCGACCT-CGCCGCGGAGCACCAGTCCGCGGCACCGC 74
   |||
   |||
   |||
97  GCTTCGCGCAGCAAGTGGCGCCGCTCATCGAGGCTCCGGGCTCTTCTACAGCGCTC 156
   |||
   |||
   |||
75  GCCTCGCGCAACAAGTGGTCCGCTCATCGAGGCTCTGGCCTCTTCTACAAGCGCTC 134
   |||
   |||
   |||
157  CAGCGCGGCGACCGTGGGAGCAGAGTCCGCTTCTATGAGCGTCTTCGCGCCACGCC 216
   |||
   |||
   |||
135  CAGTTCGGGAGCCGCGGGAGCAGAGCTCGCCTTCTACGAGCGCTTCTCCACCCACGCC 194
   |||
   |||
   |||
217  GCGTTCGCGCGCGCATCCGAGACACTCTTCTCCCGCGGTTCACGGCAGGACCTCTC 276
   |||
   |||
   |||
195  GCGTTCGCGCGCGCATCCGGACACTCTTCTCCCGGGTTTCAGCGAGCGACTCTCTC 254
   |||
   |||
   |||
277  CCACCGAGCGCAGCCCGGGAGCGGCATCTCTCACTCGTCTCTCGACACTCTCTCGCG 336
   |||
   |||
   |||
255  CCACCGAGCGCGGCGAGGGAGCGCATCCGACCTCGTCTCTCGACACTCTCTCGCG 314
   |||
   |||
   |||
337  GGGTTTCAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCATCACTGGCCACCGACT 396
   |||
   |||
   |||
315  GGCGTGGAGCGCGCTCGCTCGCGCAGATCAAGATCGGCGCATCACTGGCGCGCAGC 374
   |||
   |||
   |||
397  TCGCCGAGCGCCCTACATCGCCAAAGTGCCTCGCCCAAGGACCGCGGAGCACGAGCGTTCTG 456
   |||
   |||
   |||

```

[illegible]

| RESULT 12 | CA202412 | LOCUS | DEFINITION |
|-----------|----------|-------|------------|
|-----------|----------|-------|------------|

|                    |  |
|--------------------|--|
| 5', mRNA sequence. |  |
| CA202412           |  |
| VERSION            | CA202412.1 GI:35237651   |
| KEYWORDS           | EST.   |
| SOURCE             | Saccharum officinarum  |
| ORGANISM           | Saccharum officinarum  |
|                    | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae; PACCAD<br>clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum<br>complex.   |
| REFERENCE          | 1 (bases 1 to 671)   |
| AUTHORS            | Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.   |
| TITLE              | The libraries that made SUCSEST  |
| JOURNAL            | Genet. Mol. Biol. 24 (1-4), 1-7 (2001)   |
| COMMENT            | Contact: Arruda P<br>Centro de Biologia Molecular e Engenharia Genetica<br>Universidade Estadual de Campinas<br>Caixa Postal 6010, 13083-970, Campinas SP, Brazil<br>Tel: 55 19 3788 1137<br>Fax: 55 19 3788 1089<br>Email: parruda@unicamp.br<br>Clone distribution: clone distribution information can be found<br>through the Brazilian Clone Collection Center (BCCC) at<br><a href="http://www.bccccenter.fcav.unesp.br">http://www.bccccenter.fcav.unesp.br</a><br>Plate: 009 row: H column: 02<br>Seq primer: T7 Promoter Primer. |

## FEATURES

```

seq primer: 17, primer2: primer.
Location/Qualifiers
1. .671
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCRLFL1009H02"
/lab_host="DH10B"
/clone_lib="FL1"
/notes="Organ: Inflorescence at beginning of development (lcm-long); Vector: pSport1; Site: An unidirectional cDNA library generated from polyA+ mRNA. Inflorescence at beginning of development. cDNA was prepared from polyA+ mRNA. Plasmid System Kit (Invitrogen)."
were fractionated in a sepharose
fragments sizing between 0.8 and
directionally cloned into the vec
source of RNA and library constru
http://aucest.lad.ic.unicamp.br/v

```

## ORIGIN

|                       |        |   |       |                |
|-----------------------|--------|---|-------|----------------|
| Query Match           | 41.3%  | Score 555;  | DB 6; | Length 671;    |
| Best Local Similarity | 90.2%; | Pred. No. 1.le-119;   |       |                |
| Matches               | 605;   | Conservative  | 0;    | Mismatches 65; |
|                       |        |   |       | Indels 1;      |
|                       |        |   |       | Gaps 1;        |
| Qy                    | 172    | GGGGAGCACAGAGTGGCGCTTCTATGAGGCGTTCTCCGCCACAGCGCGCGTCCCGGCCCGC   | 231   |                |
|                       |        |   |       |                |
| Db                    | 2      | GGGGAGCACAGAGTGGCGCTTCTACGAGGCGTTCTTCCAACCAACGCGCGCGTCCCGGCCCGC | 61    |                |
|                       |        |   |       |                |
| Qy                    | 232    | ATCCGAGACACCTTCTTCCCGGGTTCACGGGACGCGACTCTCTCCGCCACCGAGGCGCAG    | 291   |                |
|                       |        |   |       |                |
| Db                    | 62     | ATCCGGGACACCTTCTTCCCGCGGTTTTCACGGCACGCGACTCTCTCCGCCACCGAGGCGCG  | 121   |                |
|                       |        |   |       |                |
| Qy                    | 292    | CCGGGGGAGCGGATCTCTACCTCTGCTCTCGAGACCTCTCTCGGGGGGTTTCAGGCGGCC    | 351   |                |
|                       |        |   |       |                |
| Db                    | 122    | CCAGGGGAGCGGATCCGCGACCTCTGCTCTCGAGACCTCTCTCGGGGGGCTGGAGGGCGCC   | 181   |                |
|                       |        |   |       |                |
| Qy                    | 352    | TGCGTCCGACAGATCAAGATCGGGGCCATCACGTGGCCACCGAGTTCGCGGGAGCCCTTAC   | 411   |                |
|                       |        |   |       |                |
| Db                    | 182    | TGCGTCCGCGACATCAAGATCGGGGCCATCACGTGGCGCGCGAGCTCGCGGAGCCCTTAC    | 241   |                |
|                       |        |   |       |                |
| Qy                    | 412    | ATCCGCCAAGTGCCCTCGGCAAGGACCGCGGGACCAAGAGCGTTCTGCTCGGATTCGCGGTC  | 471   |                |
|                       |        |   |       |                |
| Db                    | 242    | GTCGCCAAGTGCCCTCGGCAAGGACCGCGGGACCAAGAGCATTCGTCTCGGATTCGCGGTC   | 301   |                |
|                       |        |   |       |                |
| Qy                    | 472    | TCCGGCGTCCGAGTGTGTGGCGCCCGCGCGCGCGTGTGGGGGACGAGAGCGCCCGAGGTG    | 531   |                |

```
Db      302  TCCGGCGTCCGGTCTCGTCCGAGGCGCGTGTGCGGACTGAGCGCCCGAGGTG 361
Qy      532  AAGGCCATGACACCGCGCGGTCTCGCGGTCTCGCGGTCTACGTGTATCCCGTTGCC 591
Db      362  AAGGCCCTGACACCGCGCGGTCTCGCGGTCTCGCGGTCTACGTGTATCCCGTTGCC 421
Qy      592  GACGAGGGATGACTGTGCGCTCGCGCGCGGTGTACGGAGAAAAGTGGAGTCTTG 651
Db      422  GACGAGGGATGACTGTGCGCTCGCGCGCGGTGTACGGAGAAAAGTGGAGTCTTG 481
Qy      652  TCACAGCTCGCGAGCTCAAGCGGTGTTCGAGGACAGACTCTGTTCACCTTCTACTCG 711
Db      482  TCACAGCTCGCGAGCTTAAGCGGTGTTCGAGGACAGACTCTGTTCACCTTCTACTCG 541
Qy      712  GCCTCGATTTCTTGGGCTATGATGCTGCTGAGTGCAGCAGCGGAGATGGGGTGGG 771
Db      542  GCGGTGAATCTTTTGGGCTATGATGCTGCTGAGTGCAGCAGCGGAGTGAAGTGGG 601
Qy      772  GTCAAGCTGAAGCTGTGGAATTTGCCATGTGCGCGAGGTGATGGGTGATTGACCAC 831
Db      602  TTGAAGTGAA-CTTGTGACTTTTCCATTTTGTCAAAGTTGATAGGGGGAATGAACCA 660
Qy      832  AACTTCTCTGGG 842
Db      661  AATTTCTGGG 671

RESULT 13
CC724950/c
LOCUS      722 bp      DNA      linear      GSS 23-JUN-2003
DEFINITION OGLAB23TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBwa0307C22,
            genomic survey sequence.
ACCESSION  CC724950
VERSION     CC724950.1 GI:32143883
KEYWORDS   GSS.
SOURCE     Zea mays
            Zea mays
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 722)
            Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other GSSs: OGLAB23TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.

FEATURES             source
     source           1..722
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBwa0307C22"
                     /clone_lib="ZM 0.7_1.5_KB"
                     /notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"

ORIGIN
Query Match      40.2%; Score 540.8; DB 9; Length 722;
Best Local Similarity 89.8%; Pred. No. 2.4e-116;
Matches 638; Conservative 0; Mismatches 17; Indels 57; Gaps 3;

Qy      246  CTTTCCCGCGTTCCAGGACCGACTCTCTCCACCGAGGCGCAGCCCGGGAGCGCA 305
|||||
```

```
Db      722  CTTTCCCGCGTTCCAGGACCGACTCTCTCCACCGAGGCGCAGCCCGGGAGCGCA 663
Qy      306  TCCTCACCCTCGTCTCGAGACCTCTCTCGCGGGTTTCAGGCGCCCTGCGTCGACACAT 365
Db      662  TCCGACCTCGTCTCGAGACCTCTCTCGCGGGTTTCAGGCGCCCTGCGTCGACACAT 603
Qy      366  CAAAGATCGGCGCATCACGTGGCCACCGAGTTTCGCGGAGCCCTACATCGCCAAAGTGCCT 425
Db      602  CAAAGATCGGTCGACATCAGT----- 583
Qy      426  CGCCAAGGACCGCGGGACACGAGCGTTCTGCTCGGATTCGCGGTCTCGGCGTCCGAGT 485
Db      582  -----GACCACGAGCTTCTGCTCGGATTCGCGGTCTCGGCGTCCGAGT 538
Qy      486  CFTCGGCCCGGAGGCGCGGTGTGCGGACGAGAGCGCCGCGAGGTGAAGGCCATGGACAC 545
Db      537  CGTTCGCGCCCGAGGCGCGGTGTGCGGACGAGAGCGCCGCGAGGTGAAGGCTATGGACAT 478
Qy      546  CGCGCGCGTTCGCGCGCTGCTCCGCGCTACGTGTCAATCCG-TTCGCCGACGAGGGATGG 604
Db      477  TGTGCGCGTTCGCGCGCTGCTCCGCGCTACGTGTCAATCCGCTTCCGCGAGAGGGATGG 418
Qy      605  ACTGTGCGCTCGCGCGCGGTGTACGAGGAGAAAAGTGGAGTCTTGTTCACAGCTCGCGG 664
Db      417  ACTGCGCGCTCGCGCGCGGTGTACGAGGAGAAAAGTGGAGTCTTGTTCACAGCTCGCGG 358
Qy      665  AGCTCAAGCGGTGTTGAGGAGCAGACTCTGTTCCACTTCTACTCGGCGTGCATTTCTTC 724
Db      357  AGCTCAAGCGGTGTTGAGGAGGAGCAGACTCTGTTCCACTTCTACTCGGCGTGCATTTCTTC 298
Qy      725  TGGGCTATGATGCTCTCGAGTCGCGAGCGGAGATGGGGTGGGTGACGCTGAAGC 784
Db      297  TGGGCTATGATGCTCTCGAGTCGCGAGCGGAGGTGGGGTGGGTAAACAGTGAAGC 238
Qy      785  TGGTGGACTTTGCCCATGTGCGCGAGGGTGTATGGGTGATTGACCACAACTCTCTGGGCG 844
Db      237  TGGTGGACTTTGCCCATGTGCGCGAGGGTGTATGGGTGATTGACCACAACTCTCTGGGCG 178
Qy      845  GGCTCTGCTCGTGTATCAAGTTCTGTCATTTGTTCCGAGAGACTCTCTATACGCAGC 904
Db      177  GGCTCTGCTAGTGTATCAAGTTTGTCTGACATTTGTTCCAGAGACTCTCTCAGACGACG 118
Qy      905  CTTTGGGTCCTTCTTAAAGAGAGGATCCTGGCA-TTTCGATTGTATAACAAAG 955
Db      117  CTTTGGGTCCTTCTTAAAGAGAGATCCTGGCATTTTCGATTGTATAACAAAG 66

RESULT 14
BZ774817
LOCUS      593 bp      DNA      linear      GSS 14-MAR-2003
DEFINITION ii51a11.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone
            ii51a11, genomic survey sequence.
ACCESSION  BZ774817
VERSION     BZ774817.1 GI:28952260
KEYWORDS   GSS.
SOURCE     Zea mays
            Zea mays
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 593)
            Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,P., King,L., Miller,B., Muller,S., Nascento,L.,
            Zukaveni,T., McCombie,W.R. and Martienssen,R.A.
            Genomic shotgun sequences from Zea mays (methyl1-filtered)
            Unpublished (2002)
            Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
```

Plate: ii51 row: a column: 11  
Seq primer: -21M13UnivFwd  
Class: shotgun  
High quality sequence stop: 593.  
Location/Qualifiers  
1. 593

FEATURES  
source

/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ii51a11"  
/lab\_host="DH5a"  
/clone\_lib="WGS-ZmaysF (DH5a methyl filtered)"  
/notes="Organ: immature ears; Site 1: Xba I; Site 2: Xba I;  
The vector was digested with XbaI and one nucleotide was  
added by fill in the recessive 3' end. The genomic DNA  
was nebulized, end repaired, adaptor ligated and size  
fractionated using sephadex. The resulting fragments were  
between 0.8 and 1 kb and were cloned into the vector (.x/y  
reads in M13mp19, .b/g reads in pUC19). The same ligation  
was transformed into DH5a."

## ORIGIN

Query Match 39.9%; Score 535.6; DB 8; Length 593;  
Best Local Similarity 97.9%; Pred. No. 3.9e-115;  
Matches 564; Conservative 0; Mismatches 9; Indels 3; Gaps 2;  
QY 42 TCCCATATACCATGTCGACCTCCACCCGCGGAGACCAAGTCGCGGCCACCGCGCTC 101  
DB 20 TCCCATATACCATGTCGACCTCCACCCGCGGAGACCAAGTCGCGGCCACCGCGCTC 79  
QY 102 CGCAGCAAGTCGGCGCGCTCATGACGCTCGGCTCTTCTACAAAGCGCTCCAGGC 161  
DB 80 CGCAGCAAGTCGGCGCGCTCATGACGCTCGGCTCTTCTACAAAGCGCTCCAGGC 139  
QY 162 CGCGACCGTGGGAGACGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGCGCGT 221  
DB 140 CGCGACCGTGGGAGACGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACCGCGCGT 199  
QY 222 CCGGCGCGCATCGAGACACCTTTTCCCGGTTCCAGGACCGGACGCTCTCCGCCAC 281  
DB 200 CCGGCGCGCATCGAGACACCTTTTCCCGGTTCCAGGACCGGACGCTCTCCGCCAC 259  
QY 282 CGAGGCGCAGCGCGGAGCGGATCTCACTCGTCTCGACGACCTCTCGCGGGTT 341  
DB 260 CGAGGCGCAGCGCGGAGCGGATCTCGACCTCTCTCGACGACCTCTCGCGGGTT 319  
QY 342 TCAGGCGCCTGCGTCGACGACATCAAGATCGGCGCCATCACTGCGCACCGAGTTGCGC 401  
DB 320 TGAGGCGCCTGCGTCGACGACATCAAGATCGGCGCCATCACTGCGCACCGAGTTGCGC 379  
QY 402 GGAGCCCTACATCGCCAAAGTGCCTCGCAAGGACCGCGGACCAAGAGCGTTCTGCTCG 461  
DB 380 GGAGCCCTACATCGCCAAAGTGCCTCGCAAGGACCGCGGACCAAGAGCGTTCTGCTCG 439  
QY 462 ATTCCGGCTCTCCGGCTCGAGTCTCGGCGCGCGCGCGCGCGCGTGTGGCGGAGCG 521  
DB 440 ATTCCGGCTCT--TGCGTCCGAGTCTGTGGCGCGCGCGCGCGCGTGTGGCGGAGCG 497  
QY 522 CCGGAGGTGAAGCCATGACACCGCGCGGTCCGCGGTGTCTCCGGCGCTACGTGTC 581  
DB 498 CCGGAGGTGAAGGCTATGACACCGTGGGCTCCGCGGTGTCTCCGGCGCTACGTGTC 557  
QY 582 ATCCG-TTGGCGAGCGAGGATGAGTGTGGCTCG 616  
DB 558 ATCCGCTTCCGACGAGGGGATGACTTGGCGCTCG 593

RESULT 15  
CAL30686 744 bp mRNA linear EST 24-SEP-2003  
LOCUS SCCCRT1004G06.g Rt1 Saccharum officinarum cDNA clone SCCCRT1004G06  
DEFINITION 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CAL30686  
CAL30686.1 GI:35013916  
EST.  
Saccharum officinarum

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Saccharum officinarum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.  
1 (bases 1 to 744)  
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bccccenter.fcav.unesp.br  
Plate: 004 row: G column: 06  
Seq primer: T7 Promoter Primer.

FEATURES  
source

Location/Qualifiers  
1. 744  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCCRT1004G06"  
/lab\_host="DH10B"  
/clone\_lib="RT1"  
/notes="Organ: Root tips (0.3cm-long) from adult plants;  
vector: pSPori; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Root tips  
(0.3cm-long) from adult plants]. cDNA was prepared from  
polyA+ mRNA using Superscript plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

## ORIGIN

Query Match 39.1%; Score 525.6; DB 6; Length 744;  
Best Local Similarity 88.6%; Pred. No. 9e-113;  
Matches 592; Conservative 0; Mismatches 74; Indels 2; Gaps 2;  
QY 37 CCCTTCCCATACCATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGGCCACCGC 96  
DB 16 CCTCTCGCCTTCAACATGTCGACCT-CCGCCCGGAGCACCAAGTCGCGGCCACCGC 74  
QY 97 GCCTCCGCGCAGCAAGTCGGCGCGCTCATCGAGCGGCTCTTCTACAAAGCGCTC 156  
DB 75 GCCTCCGCGCAGCAAGTCGGCTCGCTCATCGAGCGGCTCTTCTACAAAGCGCTC 134  
QY 157 CAGCGCGCGGACCGTGGGAGCAGAGGTGCGCTTCTATGAGCGGTTCTCCGCCAGCGC 216  
DB 135 CAGGTCCGGGACCGCGGGAGCAGAGTCTGCGCTTCTACGAGGCGTTCTCCACCGCGC 194  
QY 217 GCGTCCGCGCGCGATCCGAGACACTTCTCCCGGTTCCACGCGCGGACTCTC 276  
DB 195 GCGTCCGCGCGCGATCCGCGGACACTTCTTCCCGGTTTCCGCGGACTCTC 254  
QY 277 CCACCGAGCGCAGCGCGGAGCGCGCATCTCTCCTCTCGACGACTCTCTCGCG 336  
DB 255 CCACCGAGCGCGCGGAGCGCGCATCCGACCTCTCTCTCGAGGACTCTCTCGCG 314  
QY 337 GGGTTTCAGCGCGCTCGGTGCGAGACATCAAGATCGCGGCCATACGTCGGCCACCGAGT 396  
DB 315 GGGTGGAGCGCGCTCGGTGCGGCGACATCAAGATCGCGGCCATACGTCGGCGCGGAGC 374

|    |     |   |     |
|----|-----|---|-----|
| Qy | 397 | TCGCCGAGCCCTACATCGCCAAAGTGCCTCGCCAGGACCGCGGGACCAAGAGCGTTCTG | 456 |
| Db | 375 | TCGCCGAGCCCTACATCGCCAAAGTGCCTCGCCAGGACCGCGGGACCAAGAGCATTCYG | 434 |
| Qy | 457 | CTCGGATTCGGGTCTCCGGCGTCCGAGTGTGCGGCCCGAGGGCGCGTGTGGGGACG    | 516 |
| Db | 435 | CTCGGATTCGGGTCTCCGGCGTCCGGTCCCGAGGGCGCGTGGGGGGACT           | 494 |
| Qy | 517 | GAGCGCCCGAGGTGAAGCCATGGACACGCGCGCGTCCGCGCGTCTCCGGCGCTAC     | 576 |
| Db | 495 | GACGCGCCGAGGTGAAGCCCTTGGACACCGCGCGTCCGCGGGGCTTCGGCGCTAC     | 554 |
| Qy | 577 | GTGTCATCCGTTGCCGACGAGGGATGGACTGTGCGCTCGCGCGCGTGTACGGAGGA    | 636 |
| Db | 555 | GTTTTATTCTGTGCCGACGAGGGATGGACTGCTCCCTCGCGCGCGGT-TACGGGGC    | 613 |
| Qy | 637 | AAAGGTGGAGTCTTGTACAGCTTCGCGAGCTCAAGCGTGTTCAGGAGCAGACTCTG    | 696 |
| Db | 614 | AAAGGGGAAACTTGGTAACATGGGCAAGCTTAAAGGGGGGTAGAGGGGAAATTTGG    | 673 |
| Qy | 697 | TTCCACTT  | 704 |
| Db | 674 | TTCCTTT   | 681 |

Search completed: June 17, 2005, 01:52:45  
Job time : 4778 secs





|    |     |  |     |
|----|-----|--|-----|
| Qy | 399 | GCCGAGCCCTACATCGCAAGTGCCTGCCAAGGACCGCGGACACAGAGGTTCTGCT      | 458 |
| Db | 634 | CCCCGTCTCGTCTCCCGTCTCTCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCT      | 693 |
| Qy | 459 | CGAATTCGCGTCTCCGCGTCCGAATGTCGCCCCCGAGGCGCGTGTGCGGACGGA       | 518 |
| Db | 694 | CGTCTCCCGTCTCGTCTCTCCCGTCTCTCGTCTCTCCCGTCTCTCGTCTCTCCCGTCTCT | 753 |
| Qy | 519 | GGCCCCGAGGTGAAGGCCATGGAACCGCCGGCGTCCGCGCGTGTCTCCGGGGTCAQT    | 578 |
| Db | 754 | CCCCGTCTCGTCTCTCCCGTCTCTCCCGTCTCTCGTCTCTCCCGTCTCTCCCGTCTCTCT | 813 |

## RESULT 2

```

US-09-130-114-2
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dama, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stab
; TITLE OF INVENTION: From Multiple Transfe
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

```

Query Match 5.2%; Score 69.6; DB 2; Length 1931;  
Best Local Similarity 45.6%; Pred. No. 7.8e-07;  
Matches 246; Conservative 0; Mismatches 294; Indels 0

|    |     |  |     |
|----|-----|--|-----|
| Qy | 39  | CCTTCCCATACATGTCGACCTTACCCGCGGAGACCAAGTCGCGGGCCACCGGC        | 98  |
| Db | 274 | CCTCGTCCTCGCCCTCCCGGCTCTCGTCTCCCGTCTCGTCTCTCCCGTCTCGT        | 333 |
| Qy | 99  | CTCGCCAGCAAGCTGGCCCGCTCATCGAGGCTCCGGCTCTTCTAAAGCCGCTCCA      | 158 |
| Db | 334 | CCTCTTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCT          | 393 |
| Qy | 159 | GGCCGGGACCGTGGGGAGCAGAGGTGCGCTTCTATGAGGGGTTCCTCGGCCACGCGC    | 218 |
| Db | 394 | CGTCTCTCTCCCGCTCTCCCGCTCTCCCGGTCTCTCGTCTCTCTCCCGCTCTCGTCT    | 453 |
| Qy | 219 | CGTCCCGGCCGCTCCGAGACACCTTCTTCCCGGTTTCCAGCGCACGCACTCTCTCC     | 278 |
| Db | 454 | CCTCCCGTCTCTCCCGTCTCTGTCCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCT    | 513 |
| Qy | 279 | CACGAGCGGACGCCGGGAGCGCATCTCTACCTCGTCTCTGACGACCTCTCTCGCGG     | 338 |
| Db | 514 | CCTCTCCCGTCTCTGTCCTCTCTCCCGTCTCTCCCGTCTCTCGTCTCTCTCCCGTCTCT  | 573 |
| Qy | 339 | GTTTCAGSGCCCTCGTCGCGAGACATCAAGATCGCGCCATCAAGTGGCCACGAGTTC    | 398 |
| Db | 574 | CCCGTCTCTCCGTCCTGTCCTCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 633 |
| Qy | 399 | GCGGAGCCCTACATCGCCAAAGTGCTCTGCCAAGGACCGCGGACACAGCGTTCCTCT    | 458 |
| Db | 634 | CCCGTCTCTGTCCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  | 693 |
| Qy | 459 | CGGATTCGCGTCTCCGSGCTCCGAGTCGTCTGGCCCCGAGGGCGCGGTGTGGCGACGA   | 518 |
| Db | 694 | GTCTCTCCCGTCTCTGTCCTCCCGTCTCTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 753 |
| Qy | 519 | CGCCCGGAGGTGAAGCCATGGACACCGCCGGGTCCGCCGCGTGTCTCCGGCGCTACGT   | 578 |
| Db | 754 | CCCGTCTCTGTCCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 813 |

### RESULT 3

```

US-09-320-878-19
; Sequence 19, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-320-878-19

```

|                       |                 |  |           |               |
|-----------------------|-----------------|--|-----------|---------------|
| Query Match           | 4.4%;           | Score 58.6;  | DB 3;     | Length 38506; |
| Best Local Similarity | 44.9%;          | Pred. No. 0.00087;   |           |               |
| Matches 223;          | Conservative 0; | Mismatches 274;  | Indels 0; | Gaps 0;       |
| Qy                    | 66              | CCGCGCGGAGCACCAGTCCGCGGGCCACCGCGCCTCCGCCACGAAGCTGGGGCCGCTCAT   | 125       |               |
| Db                    | 22151           | CCGCGGCGACGGTCTGTGGCTCTGCCCGCCCGCGCCCGGTGGCGCGGACGATGTCG       | 22210     |               |
| Qy                    | 126             | CGACGGCTCCGGCTCTTCTACAAGCGCTCCAGCGCCGCGCACCTGTGGGAGCACACGAGT   | 185       |               |
| Db                    | 22211           | CGAGGCCCTGCACGGGTCTCTCGGCTGATGACGGCTTGGCTGGCGGACGAGCGGTTCA     | 22270     |               |
| Qy                    | 186             | CGCCTTTCTATGAGGGGTTCTCCGCCCAACGCCCGCGTCCCGGCCCGCATCCGAGACACCTT | 245       |               |
| Db                    | 22271           | CGATGGCGCCTGGTGCTGTGTGACCCGCGACCGGTGCGCGCCCTTCCGGGACGGCCT      | 22330     |               |
| Qy                    | 246             | CTTCCCGCGGTTCCACGGGACCGGACTCTCTCCCAACGAGGGCGGACGCCGGGAGCCGCA   | 305       |               |
| Db                    | 22331           | GCGGTCCACCGGACAGCGCCGCGTCTGGGGCGCTTCGGCGCGTCCGCGCAGACGGAGAGCC  | 22390     |               |
| Qy                    | 306             | TCCTCACTCTGCTCTCGACGACCTCTCTCGCGGGTTTCAGGCGCCCTCGCTCGCAGACAT   | 365       |               |
| Db                    | 22391           | GGGCGGGTGTGCTCTGTCTGACCTTCGCCGGGGAAGCCCGGACGGCCGGGAGCCACCGC    | 22450     |               |
| Qy                    | 366             | CAAGATCGCGGCCATCACGTGGCCACCGAGTTTCGCGGAGCCCTACATCGCCAACTGGCT   | 425       |               |
| Db                    | 22451           | CGGGGACGGCTTGACGACCGGGAGCGCCACCGTCGCGGCGACCTCTGGAGACGCCGCCCT   | 22510     |               |
| Qy                    | 426             | CGCCAAAGNACCGGGGACCAACGAGGGTTCTGTTCGGATTTCGCGTCTCCGGGCTCCGAGT  | 485       |               |
| Db                    | 22511           | CGGACAGCGCCTCGCGACCGCCTCGGCTTCGGGCGAGCGCAGCTCGCCCTCCGGGACGG    | 22570     |               |
| Qy                    | 486             | CGTCCGCCCGGAGGGCGCGTGTGGCGNACGGAGCCCGGAGGTGAAGGCCATGGACAC      | 545       |               |
| Db                    | 22571           | GGCGCTCTCTGTACCCCGCTGGCGGGGCGCGCGGCCCGCGCGCGCTCGC              | 22630     |               |
| Qy                    | 546             | CGCCGCGCTCCGCGCG   | 562       |               |

```
Db 22631 CGCGCCGACGGCCTCG 22647
|||||
RESULT 4
US-09-141-908-1
; Sequence 1, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141.908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-141-908-1
Query Match 4.4%; Score 58.6; DB 4; Length 38506;
Best Local Similarity 44.9%; Pred. No. 0.00087;
Matches 223; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
Qy 66 CCGCGCGGACGACCAAGTCGCGGCCACCGCGCTCCGCGAGCAAGTCGGGCCGCTCAT 125
|||
Db 22151 CCGCGGACCGTCTGCTGGTGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22210
|||
Qy 126 CGAGCGCTCGGCTCTTCTACAAGCGCTTCCAGGCGCGCGCGCGCGCGCGCGCGCGCG 185
|||
Db 22211 CGAGGCGCTCGACGGTTCGCTCGGCTGATGCGAGGCTGCTGCGCGCGCGCGCGCG 22270
|||
Qy 186 CGCCTTCTATAGCGCTTCTCCGCCCAAGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCG 245
|||
Db 22271 CGATGGCGCTGCTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22330
|||
Qy 246 CTTCCCGCGGTTCCAGCGACGCGATCTCTCCCAAGCGCGCGCGCGCGCGCGCGCGCGCG 305
|||
Db 22331 CGCGTCCACGGGACAGCGCGCGCTCTGCGCGCTCTGCGCGCGCGCGCGCGCGCGCG 22390
|||
Qy 306 TCCTCACCTCGTCTCTCGACGACCTCTCGCGCGGTTTTCAGGCGCGCTCTGCGCGAGAT 365
|||
Db 22391 GGGCGGTTCTCTCTGCTGACCTCGCGGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 22450
|||
Qy 366 CAAGATCGGGCCATCACTGCGGCACCGAGTTCCGCGGAGCCCTTACATCGCCAAAGTGCCT 425
|||
Db 22451 CGGGGACGGCTGACGACCGCGGACGCGACCGTCCGCGCGCGCGCGCGCGCGCGCGCG 22510
|||
Qy 426 CGCCAAAGGACCGGGGACCGAGCGTTCTGCTCGGATTCGCGGTTCTCCGCGCTCCGAGT 485
|||
Db 22511 CGGCGCGCTCGCGACCGCGCTCGGCTCGGGGAGCGCGAGCTCGCCCTCCCGGACGG 22570
|||
Qy 486 CGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
|||
Db 22571 GGGCTCTCTGTAACCGCGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22630
|||
Qy 546 CGCGCGCGTCCGCGCG 562
|||
Db 22631 CGCGCGGACGGCCTCG 22647
|||||
RESULT 6
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
```



|    | Matches | 194;   | Conservative                   | 0;                     | Mismatches | 226; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------------------------|------------------------|------------|------|--------|----|------|----|
| Qy | 186     | CGCCTTCTATGAGCGTTCTCCGCCACGC                               | CGCGTGTC                       | CGCGCCGGCATCCGAGACACTT | 245        |      |        |    |      |    |
| Db | 2686    | CGGCTTGCGCGACCCCACTTCGGTGCCTCCGGCCGCGGCTTCGCTGCGGTGCGGCT   | 2627                           |                        |            |      |        |    |      |    |
| Qy | 246     | CTTCCCCCGGTTCCACCGCACCGACATCTCTCCCACCGAGCGCAGCCCGGGGAGCGCA | 305                            |                        |            |      |        |    |      |    |
| Db | 2626    | CCGCCCGCTTCACCGCGGGGCCCGCTCCCGCGCCCGCCCGCCCGCCCGCGCCA      | 2567                           |                        |            |      |        |    |      |    |
| Qy | 306     | TCCTCACTCGTCTCGAACACTCTCTCGGGGGTTTTAGCGGCCCTGCGTGCAGACAT   | 365                            |                        |            |      |        |    |      |    |
| Db | 2566    | GGTGGCCCTTCCCTCTCGCCCTCGGGGGAGCTCGGCACC                    | CGCGCTT                        | TAAACCCCTCCCC          | 2507       |      |        |    |      |    |
| Qy | 366     | CAGAATCGGGCCATCACTGGCCACCGAGTTTCGCCGGAGCCTTACATCGCCAA      | 425                            |                        |            |      |        |    |      |    |
| Db | 2506    | CACCGGGGTCTCTCTCGCCCCGCA                                   | CGCGCGTTCCT                    | TGCGGCGAGTTA           | 2447       |      |        |    |      |    |
| Qy | 426     | CGCCAAAGACCGGGGACCAAGAGCTTCTGCTCGGATTTCGGGTCTCCGGCGTTCGAGT | 485                            |                        |            |      |        |    |      |    |
| Db | 2446    | TGCGCGCCCGCGGAGGCCAGGGCCGCGGGGGGGTTCGCCACTAGCTGCCCGGA      | 2387                           |                        |            |      |        |    |      |    |
| Qy | 486     | CGTTCGGCCCCGAGGGCCGCTGTGGCGGA                              | CGAGCGCCCGGAGGTGAAGGCCATGGACAC | 545                    |            |      |        |    |      |    |
| Db | 2386    | GGCGGGCGCACCGAGGCCCGCCCGCCCGGACCGCGAGGGGGACCGAAA           | CGAAGAGCC                      | 2327                   |            |      |        |    |      |    |
| Qy | 546     | CGCCGGCGTTCGCCGCTGTCTCGGCGGTACGTGTATCTCGTTCGCACAGGGGGATGGA | 605                            |                        |            |      |        |    |      |    |
| Db | 2326    | CTCTCCSCAACCGGGGACGGCCCAAGTGC                              | CCCGCTT                        | CGCTGGGTGT             | 2267       |      |        |    |      |    |

RESULT 9  
 US-09-881-165-4  
 ; Sequence 4, Application US/09881165  
 ; Patent NO. 6632930  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOOD, ELIZABETH  
 ; APPLICANT: HOWARD, JOHN  
 ; APPLICANT: BAILEY, MICHELE  
 ; APPLICANT: GASTEL, FRANS VAN  
 ; APPLICANT: WANG, HUAMING  
 ; APPLICANT: WARD, MICHAEL  
 ; APPLICANT: WOODARD, SUSAN  
 ; TITLE OF INVENTION: METHOD OF INCREASING RECOVERY OF HETEROLOGOUS ACTIVE  
 ; TITLE OF INVENTION: ENZYMES PRODUCED IN PLANTS  
 ; FILE REFERENCE: 10032R  
 ; CURRENT APPLICATION NUMBER: US/09/881,165  
 ; CURRENT FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: 60/211,732  
 ; PRIOR FILING DATE: 2000-06-15  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 1082  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA encoding  
 ; OTHER INFORMATION: Organophosphate Hydrolase  
 US-09-881-165-4

|    | Query Match           | 4.3%   | Score 58;          | DB 4;           | Length 1082;      |
|----|-----------------------|--|--------------------|-----------------|-------------------|
|    | Best Local Similarity | 46.2%  | Pred. No. 0.00037; |                 |                   |
|    | Matches 230;          | Conservative   | 0;                 | Mismatches 265; | Indels 3; Gaps 1; |
| Qy | 11                    | AGTCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGACACTCCACCCGC    | 70                 |                 |                   |
| Db | 121                   | AGGCGGCTTCACTCACCCAGACATCTGGGCTCTTCGCGCGGCTCTCTCCGCG         | 180                |                 |                   |
| Qy | 71                    | CGGAGCACCAAGTCGCGCGGCCACCGCGCCTCGGCAGCAAGCTGGCGCCGCTCATCGACG | 130                |                 |                   |
| Db | 181                   | CTGGCGGGAGTCTTTCGGCTCCGCGCAAGGGCCCTGGCGAGAAGCGCTGGCGGCGCTCC  | 240                |                 |                   |

Qy 131 GCTCCGCCCTCTTCTACAGCCGCTCCAGCGCGCGACCGTGGGGAGCACGAGGTGCGCCT 190  
 Db 241 GC CGCGCCGCGCGCGCGCGTCCGACCAATCGTGGACGTGTCCACTTTCGACATCGGCC 300  
 Qy 191 TCTATGAGGCGTGTCTCCGCCACCGCGCGCTCCCGCCCGCATCCGAGACACCTTTCTTCC 250  
 Db 301 GCGACGTGTCTCTCTCGCGGAGTGTCCCGCGCGCCGACGTGCACATCGTGGCCGCCA 360  
 Qy 251 CCGGTTTCCAGGACGCGACTCTCTCCACCGAGGGGAGCCCGGGAGCCGCATCCTC 310  
 Db 361 CCGGCTCTGTGTTCAGACCGCGCCCTCTCCATCGCCTCCGCTCCGTGGAGGAGCTCACCC 420  
 Qy 311 ACCTCGTCTCGACGACCTCTCTCGCGGGTTCAGGCGCCTCGCT---CGCAGACATCA 367  
 Db 421 AGTTCTTCTCCGGAGATCCAGTAGCGCATCGAGACACCGGATCCGCGCGGCATCA 480  
 Qy 368 AGATCGCGCGCATCAGCTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAAGTGCCTCG 427  
 Db 481 TCAAGGTGGCCACACCGCGCAAGGCCACCCGTTCCAGGAGCTCGTCTCAAGGCGCGCG 540  
 Qy 428 CCAAGGACCGGGACCAAGAGGTTCTGCTCGGATTCGCGCTCTCGGCGTCCGAGTCG 487  
 Db 541 CCGCGCTCTCTTCGCCACCGCGGTCCGCGTACCACCCACCGCGCGCTCCGAGCGCG 600  
 Qy 488 TCGGCCCGGAGGCGCG 505  
 Db 601 ACGCGAGCAGAGGCGG 618

RESULT 10  
 US-09-902-540-7284  
 ; Sequence 7284, Application US/09902540  
 ; Patent No. 6833447  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Wiegand, Roger C.  
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
 ; FILE REFERENCE: 38-10(15849)B  
 ; CURRENT APPLICATION NUMBER: US/09/902,540  
 ; CURRENT FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: 60/217,883  
 ; PRIOR FILING DATE: 2000-07-10  
 ; NUMBER OF SEQ ID NOS: 16825  
 ; SEQ ID NO 7284  
 ; LENGTH: 1680  
 ; TYPE: DNA  
 ; ORGANISM: Myxococcus xanthus  
 US-09-902-540-7284

| Query Match           | 4.3%;           | Score 58;   | DB 4;     | Length 1680; |
|-----------------------|-----------------|---|-----------|--------------|
| Best Local Similarity | 48.2%;          | Pred. No. 0.00042;  |           |              |
| Matches 163;          | Conservative 0; | Mismatches 175;   | Indels 0; | Gaps 0;      |
| QY                    | 282             | CGAGGCGACGCGGGAGCGCATCTCACTCGTCTCTCGAGACCTCTCTGCGGGGTT        | 341       |              |
| DB                    | 1074            | CAAGGCCCTTCCCAAGCGCGCGCCGCTCATCTTTGCGCGCGTGGCCGCGGAGGA        | 1133      |              |
| QY                    | 342             | TCAGGCGCCTTGCGTCCGACACATCAAGATCGGGCCCATCACTGGGCCACCGAGTTGCC   | 401       |              |
| DB                    | 1134            | GCAGGGGTGCTGSGGCTCGAGTACTGTGGCGGAGACACCGCCGTGCCACGCGCGGT      | 1193      |              |
| QY                    | 402             | GGAGCCCTACATCGCCAAAGTGCTCTGCCAAGGACCGGGGACACGAGCGTTCTTGCTCGG  | 461       |              |
| DB                    | 1194            | CGCGCCAAACATCAACATCGACGCGCGCAACATCATATGGCGCACCCGCGACCTCACCGT  | 1253      |              |
| QY                    | 462             | ATTCCGGGTCTCCGGCGTCCGAGTCGTGGCGCGCGCGCGCGTGTGGCGGACGGAGCG     | 521       |              |
| DB                    | 1354            | CATTGGGCTTGGGCAAAATCCAACCTGGACGCCACACTGGTGGCCCTGGCGAAGACGCGGG | 1313      |              |
| QY                    | 522             | CCCGGAGGTGAAGGCCCATGGACACCGCGCGGTGCCCGCGTGTCTCGGGCGCTACGTGTC  | 581       |              |

Db 1314 CCGGTGGTGAAGCGGACGAGCTGTGGACCGCGCTTCTTCTACCGGTGGAGCCAGTT 1373  
QY 582 ATCCGTTCGCGAGCGAGGGATGACGTGGCTCGCG 619  
Db 1374 CAACTTCGCAAGCGGGCATTCGCCCGGTACTTCG 1411

RESULT 11

US-09-902-540-687/c  
; Sequence 687, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 687  
; LENGTH: 5054  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-687

Query Match 4.3%; Score 58; DB 4; Length 5054;  
Best Local Similarity 48.2%; Pred. No. 0.00061;  
Matches 163; Conservative 0; Mismatches 175; Indels 0; Gaps 0;  
QY 282 CGAGGCGAGCCGCGGAGCGCATCTCTCGTCTCGAGACCTCTCTCGGGGTT 341  
Db 1955 CAAAGGCCCTTCCCAAGCGCGCGCTCCATCTCTTTCGCGCGTGGCGCGAGGA 1896  
QY 342 TCAGGGCGCTGGGTGCGAGACATCAAGATCGCGCATACGTGSCCACCAGTTCCGC 401  
Db 1895 GCAGGGGCTGTGGGTTCGACTACTGTGCGGACACCCCGCTGCCACCGCGGGT 1836  
QY 402 GGAGCCCTACATGCCAAGTGCTCTCCCAAGGACCGCGGACACAGAGCTTCTGCTCGG 461  
Db 1835 CGCGGCCAATCAACATCATGAGGGGCCAATCATCATGSCGACCCGCGACCTCACGT 1776  
QY 462 ATTCCGCTCTCGGGCTCGAGTGTGTGCGCCCGAGGGCGGCTGTGCGGACGAGCG 521  
Db 1775 CATTCGCTTGGGCAATTCACCTTGAGCGCCACACTGTGTGCGCTTGGCGAAGACGCGG 1716  
QY 522 CCGGAGGTGAAGGCCATGACACCGCGCGTTCGCGCGCTCCGCGCTACGTGTC 581  
Db 1715 CCGGTGGTGAAGCGGACCACTGTGCGACCGCGCTTCTTCTACCGGTGGACCAAGTT 1656  
QY 582 ATCCGTTCGCGAGCGGGGATGACGTGTGCGCTCGCG 619  
Db 1655 CAACTTCGCAAGCGGGCATTCGCCCGCGGTACTTCG 1618

RESULT 12

US-09-799-451-881  
; Sequence 881, Application US/09799451  
; Patent No. 6783969  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yundong

; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Ghosh, Reena  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 803  
; CURRENT APPLICATION NUMBER: US/09/799,451  
; CURRENT FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 948  
; SOFTWARE: pc\_FL\_genes Version 2.0  
; SEQ ID NO 881  
; LENGTH: 4530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (330)..(4265)  
US-09-799-451-881

Query Match 4.3%; Score 57.6; DB 4; Length 4530;  
Best Local Similarity 46.7%; Pred. No. 0.00074;  
Matches 183; Conservative 0; Mismatches 209; Indels 0; Gaps 0;  
QY 140 TTTTCTACAAGCCGCTCCAGGCGCGCGCGTGGGGAGCAGAGGTGCGCTTCTATGAGG 199  
Db 1384 TCTTCGTGTGGGACCGAGGCAGAGGCTTGTGGGAAGGCAAGGTCTCGGTGGAGGTGG 1443  
QY 200 GGTTCCTCCGCCACCGCGCGCTCCCGCGCGCATCCGAGACACCTTCTTCCCGCGTTC 259  
Db 1444 TGACCGCTGGTGGAGCCGCTCCCGCGCGCTCTGCTCGCGCGCGACTCGGAGATCG 1503  
QY 260 ACGGCACCGACTCTCTCCACAGAGGCGCAGCCGCGGAGCGCATCTCACTCGTCC 319  
Db 1504 AGAAGGGAGATGCTTCAGCGGAGGAGGCCAGGCTGGCGTGTCTCTTCCGCG 1563  
QY 320 TCGACGACTCTCTCGGGGTTTTCAGGCGCTTCGCTGCGAGACATCAAGATCGGCGCA 379  
Db 1564 CGGGCGCGCGCGCGCTTACACCGCGCTTCGCGCACCCACCGCCAGCGCC 1623  
QY 380 TCAGTGGCCACCGAGTTTCGCGGAGGCCCTACATCGCAAGTGCCTCGCAAGGACCGG 439  
Db 1624 TTCTTCAGCTTCCGCTCCGCGCGCGCGCGAGGGGAGCGACTTCTTGTCTCTGATGCGG 1683  
QY 440 GGACCACGAGGCTTCTGCTCGGATTCGCGGTTCCGCGGTCCGAGTCTCGCGCCCGGAGG 499  
Db 1684 AGTCGAGCGCGAGGCGCGCTTCGAGTGGACCTAGGGAGCGCGCTCCGCGCCCGCGG 1743  
QY 500 CGCGCGTGTGGCGAGCGCGCGCGCGAGGTG 531  
Db 1744 CGCGCGACTCGCGTGGCGCGCGCTGGACCTG 1775

RESULT 13

US-09-773-816-1  
; Sequence 1, Application US/09773816  
; Patent No. 6340774  
; GENERAL INFORMATION:  
; APPLICANT: Stanford University  
; APPLICANT: Khosla, Chaitan  
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR  
; TITLE OF INVENTION: ANTAGONISTS  
; FILE REFERENCE: 28600-20210.00  
; CURRENT APPLICATION NUMBER: US/09/773,816  
; CURRENT FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 60/243,458  
; PRIOR FILING DATE: 2000-10-25  
; PRIOR APPLICATION NUMBER: US 60/179,305  
; PRIOR FILING DATE: 2000-01-31

```

; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23673
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(23623)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1

Query Match
Best Local Similarity 4.3%; Score 57.6; DB 3; Length 23673;
Matches 181; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 311 ACCTCGTCTCGACGACCTCTCTCGCGGGTTTCAGGCGCCCTCGTGCAGACATCAAGA 370
Db 14307 ACCCGGGCACTGCTGCGGGTGTTCGCGCTGCGCGCTCTTCGGGATGCTGCGATGC 14366

QY 371 TCGGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAAGTGCCTGCCA 430
Db 14367 GCGGCGACATCCGCGCCCAACCGGGGTGCTGAAGACCAACGACATGGTACGCTGGCG 14426

QY 431 AGGACCGGGGACACAGAGGTTCTGCTCGGATTCGCGCTCTCGGGCTCCGAGTGTGCG 490
Db 14427 CGTCTGCGCCGACGGGCGGACCTCGACCTGTCTGACGTGTCTGACGTGTCTGGCCCG 14486

QY 491 GCCCGGAGGCGCGTGTGCGGACGAGCGCCCGGAGGTGAAGGCCATGACACCGCGCG 550
Db 14487 CCGACAAGGCGCGCGTGTGCGCTGACCTGCGCGCGCGCGCGGAAATGCTGGGCG 14546

QY 551 GCGTCGCGCGCTGCTCCGCGCTACGTGTCA---TCCGTTGCGGAGGAGGATGACT 607
Db 14547 GCGTCGCGTGGGGCTGTGCTGCGACGACCGCTGTGCCGCTGCGAGCGGAGTGGCG 14606

QY 608 GTGCGTCCGCGCGGCGGTGTACGAGGAAAGTGAAGTGTGTCACAGCTCGCGGAGC 667
Db 14607 ACGTGTGCGCGCGCGGTGTGACCGCTCAAGGCGACCGCGCTGCGACGCTGGCGGACG 14666

QY 668 TCAAGGCG 675
Db 14667 TCAGCGG 14674

RESULT 14
US-09-105-537-32
; Sequence 32, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 11220
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-32

Query Match
Best Local Similarity 4.2%; Score 57; DB 3; Length 11220;
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 66 CCGCGCGGAGCACCAAGTCCGCGCCACCGCGCTCCGCGGAGCAAGTGGCGCGCTCAT 125
Db 24009 CCGCGCGGAGCGTCTCTGTGSCCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 24068

QY 126 CGACGGTCCGGCTCTTCTACAGCGCTCCAGGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 24069 CGAGGCGCTGACGGGTGCTGCTGATGACGCGCTGCTGCGCGCGCGCGCGCGCG 24128

QY 186 CGCTTCTATGAGCGCTTCTCCGCCACCGCGCTCCCGCGCGCGCGCGCGCGCGCGCT 245
Db 24129 CGATGGCGCGCTGCTGCTGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 24188

QY 246 CTTTCCCGCGGTTCACGGCACGCGACTCTCTCCACCGAGCGCGCGCGCGCGCGCGCG 305
Db 8322 CCGCGCGGACCGTCTGTGTCGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8381

```

```

QY 126 CGACGGCTCCGGCTCTTCTACAGCGCTCCAGGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 8382 CGAGGCGCTGACGGGTGCTGCTGATGACGCGCTGCTGCGCGCGCGCGCGCGCGCG 8441

QY 186 CGCTTCTATGAGCGGTTCCTCCGCCACCGCGCGCTCCCGCGCGCGCGCGCGCGCGCT 245
Db 8442 CGATGGCGCGCTGCTGCTGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 8501

QY 246 CTTTCCCGCGGTTCACGGCACGCGACTCTCTCCACCGAGGCGCGCGCGCGCGCGCGCG 305
Db 8502 GCGGTTCACGGCACGCGCGCTCTGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCG 8561

QY 306 TCTTACCTCTGCTCTCGACGACCTCTCGCGGGTTTCAGGCGCGCGCTGCTGCGCGACAT 365
Db 8562 GGGCGGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8621

QY 366 CAAGATCGCGCCATCAGTGGCCACCGAGTTCGCGCGGAGCCCTACATTCGCCAAGTGCCT 425
Db 8622 CGGGGACGCGCTGACGACCGGGGACCGACCGTCCGCGCGCGCGCGCGCGCGCGCT 8681

QY 426 CGCAAGGACCGGGGACCGAGCGTTCGCTCGGATTCGCGCTCTCCGCGCTCTCCGCGCT 485
Db 8682 CGGACGCGCTCTGCGACCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 8741

QY 486 CGTCCGCGCGCGCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
Db 8742 GGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8801

QY 546 CGCGCGCGCTCCGCGCG 562
Db 8802 CGCGCGCGCGCGCGCTCG 8818

```

```

RESULT 15
US-09-105-537-5
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

```

```

Query Match
Best Local Similarity 4.2%; Score 57; DB 3; Length 36778;
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 66 CCGCGCGGAGCACCAAGTCCGCGCCACCGCGCTCCGCGGAGCAAGTGGCGCGCTCAT 125
Db 24009 CCGCGCGGAGCGTCTCTGTGSCCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 24068

QY 126 CGACGGTCCGGCTCTTCTACAGCGCTCCAGGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 24069 CGAGGCGCTGACGGGTGCTGCTGATGACGCGCTGCTGCGCGCGCGCGCGCGCG 24128

QY 186 CGCTTCTATGAGCGCTTCTCCGCCACCGCGCTCCCGCGCGCGCGCGCGCGCGCT 245
Db 24129 CGATGGCGCGCTGCTGCTGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 24188

QY 246 CTTTCCCGCGGTTCACGGCACGCGACTCTCTCCACCGAGCGCGCGCGCGCGCGCGCG 305
Db 8322 CCGCGCGGACCGTCTGTGTCGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8381

```



|    |       |    |      |      |        |        |      |       |      |      |       |      |       |       |     |      |       |      |      |      |    |      |     |       |
|----|-------|----|------|------|--------|--------|------|-------|------|------|-------|------|-------|-------|-----|------|-------|------|------|------|----|------|-----|-------|
| Db | 24189 | GC | GTCC | ACGG | GACAGG | CCGCCG | CTCG | GGGC  | CTCG | GC   | GAGAG | CCC  | 24248 |       |     |      |       |      |      |      |    |      |     |       |
| Qy | 306   | TC | CTAC | CTCG | TCTCG  | AGAC   | CTCT | CG    | GGG  | TTTC | AG    | GC   | CTCG  | TCTCG | CAG | ACAT | 365   |      |      |      |    |      |     |       |
| Db | 24249 | GG | CCCG | GTTC | GTCT   | CTCG   | ACCT | CG    | CCCG | GAAG | CCCG  | AG   | CCCG  | GGG   | AG  | CCG  | 24308 |      |      |      |    |      |     |       |
| Qy | 366   | CA | AGAT | CG   | GGCC   | ATC    | AC   | CGT   | GG   | CCAC | CG    | AGTT | CG    | CCG   | AG  | CCCT | AC    | ATCG | CCAA | GTGC | CT | 425  |     |       |
| Db | 24309 | CG | GGG  | AGG  | CCCT   | GAC    | CA   | CG    | GGG  | AC   | CG    | CA   | CC    | CTCG  | GG  | GGC  | AC    | CTCT | GG   | AG   | AG | CGCC | CT  | 24368 |
| Qy | 426   | CG | CCAA | GG   | ACCG   | CGG    | AC   | CA    | CG   | AG   | CGTT  | CT   | GT    | CG    | GA  | TTCC | GG    | CTCG | GG   | GTCC | GA | GT   | 485 |       |
| Db | 24369 | CG | GC   | AG   | CG     | CTCG   | CA   | CG    | CC   | CTCG | GG    | CTCG | GG    | CTCG  | GG  | CTCG | GG    | CTCG | GG   | CTCG | GG | CTCG | GG  | 24428 |
| Qy | 486   | CG | TC   | GG   | CCCG   | AGG    | CG   | CGT   | GT   | GG   | CG    | AG   | CG    | CG    | CG  | AG   | CG    | CG   | CG   | CG   | CG | CG   | CG  | 545   |
| Db | 24429 | GG | CG   | CTCT | CGT    | AC     | CC   | CG    | CT   | GG   | CG    | GG   | CG    | CG    | CG  | CG   | CG    | CG   | CG   | CG   | CG | CG   | CG  | 24488 |
| Qy | 546   | CG | CCG  | CG   | GTCC   | CG     | CG   | CG    | 562  |      |       |      |       |       |     |      |       |      |      |      |    |      |     |       |
| Db | 24489 | CG | GG   | CCG  | AG     | CG     | CTCG | 24505 |      |      |       |      |       |       |     |      |       |      |      |      |    |      |     |       |

Search completed: June 17, 2005, 01:57:11  
Job time : 267 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 21:47:21 ; Search time 786 Seconds  
(without alignments)  
10122.306 Million cell updates/sec

Title: US-10-042-894A-7  
Perfect score: 1344  
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID          | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 1344  | 100.0       | 1344   | 6 AAD43514  | Aad43514 Maize ino |
| 2          | 864.4 | 64.3        | 922    | 6 AAD43513  | Aad43513 Maize ino |
| 3          | 844.8 | 62.9        | 1169   | 6 AAD43511  | Aad43511 Maize ino |
| 4          | 837.4 | 62.3        | 923    | 6 AAD43512  | Aad43512 Maize ino |
| 5          | 740.4 | 55.1        | 3416   | 6 AAD43522  | Aad43522 Maize ino |
| 6          | 696   | 51.0        | 899    | 6 AAD43518  | Aad43518 Maize ino |
| 7          | 470.4 | 35.0        | 643    | 6 AAD43519  | Aad43519 Maize ino |
| 8          | 298.4 | 22.2        | 519    | 6 AAD43520  | Aad43520 Maize ino |
| 9          | 191.6 | 14.3        | 353    | 6 AAD43521  | Aad43521 Maize ino |
| 10         | 184.6 | 13.7        | 1195   | 6 AAD43516  | Aad43516 Eucalyptu |
| 11         | 179   | 13.3        | 1105   | 6 AAD43515  | Aad43515 Soybean 1 |
| 12         | 149.8 | 11.1        | 1020   | 6 AAD43517  | Aad43517 P. argent |
| 13         | 148.6 | 11.1        | 1168   | 3 AAC39023  | Aac39023 Arabidops |
| 14         | 147   | 10.9        | 1104   | 3 AAC33685  | Aac33685 Arabidops |
| 15         | 147   | 10.9        | 1243   | 3 AAC38692  | Aac38692 Arabidops |
| 16         | 145.6 | 10.8        | 1130   | 3 AAC48750  | Aac48750 Arabidops |
| 17         | 142   | 10.6        | 1130   | 3 AAC40313  | Aac40313 Arabidops |
| 18         | 104.6 | 7.8         | 464    | 6 ABL93575  | Ab193575 Arabidops |
| 19         | 72.4  | 5.4         | 113193 | 8 AAD54645  | Aad54645 Streptomy |
| 20         | 71.4  | 5.3         | 135638 | 10 ABX34289 | Abx34289 S. atrool |

|    |      |     |        |             |                    |
|----|------|-----|--------|-------------|--------------------|
| 21 | 68.6 | 5.1 | 1383   | 10 ADG33751 | Adg33751 Actinomyc |
| 22 | 68.6 | 5.1 | 89421  | 6 AAL40781  | Aal40781 8842int 9 |
| 23 | 65.4 | 4.9 | 484    | 10 ADE82058 | Ade82058 Arabidops |
| 24 | 65   | 4.8 | 897    | 8 ADA71279  | Ada71279 Rice gene |
| 25 | 63.8 | 4.7 | 5452   | 10 ADC86736 | Adc86736 Human GPC |
| 26 | 63.4 | 4.7 | 3849   | 4 AAF25795  | Aaf25795 S. chryso |
| 27 | 63.4 | 4.7 | 9975   | 8 AAL61173  | Aal61173 Actinosyn |
| 28 | 63.4 | 4.7 | 82746  | 8 AAL61224  | Aal61224 Actinosyn |
| 29 | 63.2 | 4.7 | 1227   | 6 ABL61294  | Ab161294 N. unifor |
| 30 | 63.2 | 4.7 | 24379  | 2 AAT93095  | Aat93095 Streptomy |
| 31 | 63.2 | 4.7 | 24379  | 2 AAV25925  | Aav25925 Streptomy |
| 32 | 62.6 | 4.7 | 11619  | 13 ADQ91715 | Adq91715 Polyketid |
| 33 | 62.6 | 4.7 | 164051 | 13 ADQ91695 | Adq91695 Polyketid |
| 34 | 62.2 | 4.6 | 1084   | 10 ADB78935 | Adb78935 Rice tran |
| 35 | 62   | 4.6 | 2000   | 8 ADA71938  | Ada71938 Rice gene |
| 36 | 62   | 4.6 | 5877   | 6 ABS78681  | Abs78681 Kitasatos |
| 37 | 62   | 4.6 | 27541  | 4 AAD17185  | Aad17185 Streptomy |
| 38 | 62   | 4.6 | 125401 | 4 AAD17186  | Aad17186 Streptomy |
| 39 | 61.8 | 4.6 | 7788   | 10 ADI23894 | Adi23894 Streptomy |
| 40 | 61.8 | 4.6 | 37360  | 10 ADI23892 | Adi23892 Streptomy |
| 41 | 61.2 | 4.6 | 2475   | 12 ADJ44598 | Adj44598 Plant cDN |
| 42 | 61.2 | 4.6 | 3354   | 8 AAL61181  | Aal61181 Actinosyn |
| 43 | 61   | 4.5 | 2466   | 13 ADT43880 | Adt43880 Bacterial |
| 44 | 61   | 4.5 | 53789  | 2 AAV21187  | Aav21187 Amycolato |
| 45 | 61   | 4.5 | 64492  | 12 ADK16023 | Adk16023 Streptomy |

ALIGNMENTS

RESULT 1  
AAD43514  
ID AAD43514 standard; DNA; 1344 BP.  
XX  
AC AAD43514;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 52..921  
FT /\*tag= a  
FT /product= "Maize IPPK protein #4"

PN WO200259324-A2.  
XX  
XX PD 01-AUG-2002.  
XX  
XX PD 09-JAN-2002; 2002WO-US003120.  
XX  
XX PD 12-JAN-2001; 2001US-0261465P.  
XX  
XX PD (PTON-) PIONEER HI-BRED INT INC.  
XX  
XX PD Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
XX  
XX PD WPI; 2002-636540/68.  
XX  
XX PD P-PSDB; AAE26196.

PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.

XX Claim 1; Page 64-65; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)

CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA

Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other; ;

| Query Match           | 100.0%         | Score 1344   | DB 6     | Length 1344 |
|-----------------------|----------------|--|----------|-------------|
| Best Local Similarity | 100.0%         | Prod. No. 2.3e-271   |          |             |
| Matches 1344          | Conservative 0 | Mismatches 0   | Indels 0 | Gaps 0      |
| Qy                    | 1              | GCAGAGTCAGTCCGGTCAACCCCTCGCGCCCATAGTCCCTTCCCACATACCATGTCGGAC | 60       |             |
| Db                    | 1              | GCACGAGTCAGTCCGTCACCCCTCGCGCCCATAGTCCCTTCCCACATACCATGTCGGAC  | 60       |             |
| Qy                    | 61             | CTCCACCCGCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGCTGGGCCG     | 120      |             |
| Db                    | 61             | CTCCACCCGCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGCTGGGCCG     | 120      |             |
| Qy                    | 121            | CTCATCGACGGCTCGGGCTCTTCTACAGCGCGTCCAGGCGCGGACCGTGGGAGGAC     | 180      |             |
| Db                    | 121            | CTCATCGACGGCTCGGGCTCTTCTACAGCGCGTCCAGGCGCGGACCGTGGGAGGAC     | 180      |             |
| Qy                    | 181            | GAGTGCCTCTATAGAGCGTTCTCCGCCACCGCGCGCTCCGGGCCCGCATCCGAGAC     | 240      |             |
| Db                    | 181            | GAGTGCCTCTATAGAGCGTTCTCCGCCACCGCGCGCTCCGGGCCCGCATCCGAGAC     | 240      |             |
| Qy                    | 241            | ACCTTCCTCCCGGTTCCACGGGACCGGCACTCTCTCCCAACCGAGGCGCAGCCGGGAG   | 300      |             |
| Db                    | 241            | ACCTTCCTCCCGGTTCCACGGGACCGGCACTCTCTCCCAACCGAGGCGCAGCCGGGAG   | 300      |             |
| Qy                    | 301            | CCGCATCTCTACCTCTCGACGACTCTCCGGGGTTTCAGGCGCCCTGGTGCACA        | 360      |             |
| Db                    | 301            | CCGCATCTCTACCTCTCGACGACTCTCCGGGGTTTCAGGCGCCCTGGTGCACA        | 360      |             |
| Qy                    | 361            | GACATCAAGATCGCGCCCATCACGTGGCCACCGAGTTCGCGGAGCCCTACATGCCCAAG  | 420      |             |
| Db                    | 361            | GACATCAAGATCGCGCCCATCACGTGGCCACCGAGTTCGCGGAGCCCTACATGCCCAAG  | 420      |             |
| Qy                    | 421            | TGCTTCGCAAGACCGGGGACACAGCGTTCTGTCTCGGATTTCGGCTTCGGCGCTC      | 480      |             |
| Db                    | 421            | TGCTTCGCAAGACCGGGGACACAGCGTTCTGTCTCGGATTTCGGCTTCGGCGCTC      | 480      |             |
| Qy                    | 481            | CGAGTCGTGCGCCCGGAGCGCGTGTGGCGGACCGGAGCCCGGAGGTCGAAGGCCATG    | 540      |             |
| Db                    | 481            | CGAGTCGTGCGCCCGGAGCGCGTGTGGCGGACCGGAGCCCGGAGGTCGAAGGCCATG    | 540      |             |
| Qy                    | 541            | GACACCGCGCGTCCGCGCGGTCTCGCGGCTACGTGTTCATCCGTTTCCGACGAGGGG    | 600      |             |
| Db                    | 541            | GACACCGCGCGTCCGCGCGGTCTCGCGGCTACGTGTTCATCCGTTTCCGACGAGGGG    | 600      |             |
| Qy                    | 601            | ATGACCTGTGCTCCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTTGTCAAGCTG       | 660      |             |
| Db                    | 601            | ATGACCTGTGCTCCGCGCGCGGTGTACGGAGGAAAGGTGAGTCTTGTCAAGCTG       | 660      |             |
| Qy                    | 661            | CGCGAGCTCAAGGGGTGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACGCGGTCAATT  | 720      |             |
| Db                    | 661            | CGCGAGCTCAAGGGGTGTGTTCGAGGAGCAGACTCTGTTCCACTTCTACGCGGTCAATT  | 720      |             |
| Qy                    | 721            | CTTCTGGGCTATGATGTCTGCTCAGTCGACGAGCGGAGATGGGGTGGGTCAGGTG      | 780      |             |
| Db                    | 721            | CTTCTGGGCTATGATGTCTGCTCAGTCGACGAGCGGAGATGGGGTGGGTCAGGTG      | 780      |             |
| Qy                    | 781            | AAGCTGGTGGACTTTGCCCATCTCGCCGAGGGTGTGGGTGATTCACCAACTTCCTG     | 840      |             |
| Db                    | 781            | AAGCTGGTGGACTTTGCCCATCTCGCCGAGGGTGTGGGTGATTCACCAACTTCCTG     | 840      |             |
| Qy                    | 841            | GGCGGGCTCTGCTCGCTGATCAAGTTTCGTTCTGACATTTGTTCCGAGACTCTCATACG  | 900      |             |

|    |   |  |
|----|---|--|
| PT | animal feed.  |  |
| XX |   |  |
| PS | Claim 1; Page 62-63; 86pp; English.                                       |  |
| XX |   |  |
| CC | The invention relates to novel inositol polyphosphate kinase (IPPK)       |  |
| CC | polypeptides and polynucleotides. Sequences of the invention are useful   |  |
| CC | in modulating the phytic acid biosynthesis by decreasing phytate and/or   |  |
| CC | increasing non-phytate phosphorous to improve the nutritional value of    |  |
| CC | animal feed, or to reduce the environmental impact of animal waste.       |  |
| CC | Polynucleotides of the invention are to produce transgenic plants with an |  |
| CC | altered phenotype. IPPK proteins are used to screen compounds that        |  |
| CC | modulate their activity and raising anti-idiotypic antibodies. The        |  |
| CC | present sequence is maize IPPK DNA  |  |
| XX |   |  |
| SQ | Sequence 922 BP; 137 A; 327 C; 288 G; 170 T; 0 U; 0 Other;                |  |
|    | Query Match 64.3%; Score 864.4; DB 6; Length 922;                         |  |
|    | Best Local Similarity 98.8%; Pred. No. 4.2e-171;                          |  |
|    | Matches 869; Conservative 0; Mismatches 11; Indels 0; Gaps 0;             |  |
| QY | 42 TCCCATACCATGTCGAGCTCCACCCGCGGAGCACCAGTCGCGGCCACCGCGCTC 101             |  |
| DB | 43 TCCCATACCATGTCGAGCTCCACCCGCGGAGCACCAGTCGCGGCCACCGCGCTC 102             |  |
| QY | 102 CGCCAGCAAGTGGGCGGCTCATCGAGGCTTCGCGCTCTTCTACAAGCGCTCCAGGC 161          |  |
| DB | 103 CGCCAGCAAGCGGCGGCTCATCGAGGCTTCGCGCTCTTCTACAAGCGCTCCAGGC 162           |  |
| QY | 162 CGGCGACGCTGGGAGCAGAGTGCCTTCTATGAGGCTTCGCGGCCACCGCGCT 221              |  |
| DB | 163 CGGCGACGCTGGGAGCAGAGTGCCTTCTATGAGGCTTCGCGGCCACCGCGCT 222              |  |
| QY | 222 CCGCGCCGATCCGAGACACCTTCTCCCGGTTTCAACGCGCAGCGACTCTCCCGAC 281           |  |
| DB | 223 CCGCGCCGATCCGAGACACCTTCTCCCGGTTTCAACGCGCAGCGACTCTCCCGAC 282           |  |
| QY | 282 CGAGCGCGAGCCGCGGAGCGCATCTCTCCTGCTTCGACGACTCTCTCGCGGTT 341             |  |
| DB | 283 CGAGCGCGAGCCGCGGAGCGCATCTCTCCTGCTTCGACGACTCTCTCGCGGTT 342             |  |
| QY | 342 TCAGGCGCTGCTGCGAGACATCAAGTCCGCGGCATCAGTGGCCACCGAGTTCGC 401            |  |
| DB | 343 TCAGGCGCTGCTGCGAGACATCAAGTCCGCGGCATCAGTGGCCACCGAGTTCGC 402            |  |
| QY | 402 CGAGCGCTTACATCGGCAAGTCTCGCCAGCAGCGGACCGAGGTTCTGCTCG 461               |  |
| DB | 403 CGAGCGCTTACATCGGCAAGTCTCGCCATGAGCAGCGGACCGAGGTTCTGCTCG 462            |  |
| QY | 462 ATTCGCGCTCTCGGCGTCCGAGTCTGTCGCGCCCGAGGCGCGTGTGGCGAGCG 521             |  |
| DB | 463 ATTCGCGCTCTCGGCGTCCGAGTCTGTCGCGCCCGAGGCGCGTGTGGCGAGCG 522             |  |
| QY | 522 CCGGAGTGAAGGCGATGGAACCGCGCGTTCGCGCGTCTCGCGCGCTAGTGTTC 581             |  |
| DB | 523 CCGGAGTGAAGGCGATGGAACCGCGCGTTCGCGCGTCTCGCGCGCTAGTGTTC 582             |  |
| QY | 582 ATTCGCTTCGCGAGCGGATGAGTGTGCTTCGCGCGGCTGTACGAGGAAAGG 641               |  |
| DB | 583 ATTCGCTTCGCGAGCGGATGAGTGTGCTTCGCGCGGCTGTACGAGGAAAGG 642               |  |
| QY | 642 TGGAGTCTTGTACAGCTCGCGAGCTCAAGCGTGTGTTTCGAGGAGCAGACTCTGTCCA 701        |  |
| DB | 643 TGGAGTCTTGTACAGCTCGCGAGCTCAAGCGTGTGTTTCGAGGAGCAGACTCTGTCCA 702        |  |
| QY | 702 CTTTACTCGGCGTCAATTTCTTCTGGGCTATGATGCTGTGAGTGCAGAGCGGAG 761            |  |
| DB | 703 CTTTACTCGGCGTCAATTTCTTCTGGGCTATGATGCTGTGAGTGCAGAGCGGAG 762            |  |
| QY | 762 TGGGGTGGGGTCAAGCTGAGTGTGACTTTCCTATGCTGGCCGAGGGTGTGGGGT 821            |  |
| DB | 763 TGGGGTGGGGTCAAGCTGAGTGTGACTTTCCTATGCTGGCCGAGGGTGTGGGGT 822            |  |
| QY | 822 GATTGACCAACTTCTCTGGGGGGCTCTGCTCGCTGATCAAGTCTGTTTCTGACATTGT 881        |  |
| DB |   |  |
| DB | 923 GATTGACCAACTTCTCTGGGGGGCTCTGCTCGCTGATCAAGTCTGTTTCTGACATTGT 882        |  |
| QY | 882 TCCGAGAGCTCTCATAGCAGCGCTTTTGGGTCTCTTTTAA 921                          |  |
| DB | 883 TCCAGAGACTCTCATAGCAGCGCTTTTGGGTCTCTTTAA 922                           |  |
|    | RESULT 3  |  |
| ID | AAD43511 standard; DNA; 1169 BP.  |  |
| XX | AAD43511  |  |
| AC | AAD43511;   |  |
| XX |   |  |
| DT | 14-NOV-2002 (first entry)   |  |
| XX |   |  |
| DE | Maize inositol polyphosphate kinase (IPPK) DNA #1.                        |  |
| XX |   |  |
| KW | Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;          |  |
| KW | nutritional value; animal feed; transgenic; gene; ds.                     |  |
| OS | Zea mays.   |  |
| XX |   |  |
| PH | Location/Qualifiers   |  |
| FT | Key 84..806   |  |
| FT | CDS /*tag= a  |  |
| FT | /product= "Maize IPPK protein #1"   |  |
| XX |   |  |
| PN | WO200259324-A2.   |  |
| XX |   |  |
| PD | 01-AUG-2002.  |  |
| XX |   |  |
| PF | 09-JAN-2002; 2002WO-US003120.   |  |
| XX |   |  |
| PR | 12-JAN-2001; 2001US-0261465P.   |  |
| XX |   |  |
| PA | (PTON-) PIONEER HI-BRED INT INC.  |  |
| XX |   |  |
| PI | Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;                          |  |
| XX |   |  |
| DR | WPI; 2002-636540/68.  |  |
| DR | P-PSDB; AAE26193.   |  |
| XX |   |  |
| PT | New inositol polyphosphate kinase polynucleotides and polypeptides,       |  |
| PT | useful in modulating phytic acid biosynthesis by decreasing phytate or    |  |
| PT | increasing non-phytate phosphorous to improve the nutritional value of    |  |
| PT | animal feed.  |  |
| XX |   |  |
| PS | Claim 1; Page 58-59; 86pp; English.                                       |  |
| XX |   |  |
| CC | The invention relates to novel inositol polyphosphate kinase (IPPK)       |  |
| CC | polypeptides and polynucleotides. Sequences of the invention are useful   |  |
| CC | in modulating the phytic acid biosynthesis by decreasing phytate and/or   |  |
| CC | increasing non-phytate phosphorous to improve the nutritional value of    |  |
| CC | animal feed, or to reduce the environmental impact of animal waste.       |  |
| CC | Polynucleotides of the invention are to produce transgenic plants with an |  |
| CC | altered phenotype. IPPK proteins are used to screen compounds that        |  |
| CC | modulate their activity and raising anti-idiotypic antibodies. The        |  |
| CC | present sequence is maize IPPK DNA  |  |
| XX |   |  |
| SQ | Sequence 1169 BP; 238 A; 362 C; 347 G; 222 T; 0 U; 0 Other;               |  |
|    | Query Match 62.9%; Score 844.8; DB 6; Length 1169;                        |  |
|    | Best Local Similarity 97.2%; Pred. No. 5.5e-167;                          |  |
|    | Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;             |  |
| QY | 42 TCCCATACCATGTCGAGCTCCACCCGCGGAGCACCAGTCGCGGCCACCGCGCTC 101             |  |
| DB | 74 TCCCATACCATGTCGAGCTCCACCCGCGGAGCACCAGTCGCGGCCACCGCGCTC 133             |  |
| QY | 102 CGCCAGCAAGTGGGCGGCTCATCGAGGCTTCGCGCTCTTCTACAAGCGCTCCAGGC 161          |  |
| DB | 134 CGCCAGCAAGTGGGCGGCTCATCGAGGCTTCGCGCTCTTCTACAAGCGCTCCAGGC 193          |  |

QY 162 CGGGACCGTGGGAGACGAGGTGCGCTTCTATGAGGCGTTCTCCGCCACGCGCGCT 221  
 Db 194 CGGGACCGTGGGAGACGAGGTGCGCTTCTATGAGGCGTTCTCCGCCACGCGCGCT 253  
 QY 222 CCGGCGCGCATCCGAGACACCTTCTCCCGGTTCCACGGGACGCGACTCTCCGCCAC 281  
 Db 254 CCGGCGCGCATCCGAGACACCTTCTCCCGGTTCCACGGGACGCGACTCTCCGCCAC 313  
 QY 282 CGAGCGCGAGCCCGGGAGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 341  
 Db 314 CGAGCGCGAGCCCGGGAGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 373  
 QY 342 TCAGCGCGCTCGCTCGAGACATCAAGATCGGCGCATCACTGCGGCGACCGAGTTCGCC 401  
 Db 374 TGAGCGCGCTCGCTCGAGACATCAAGATCGGCGCATCACTGCGGCGACCGAGTTCGCC 433  
 QY 402 GGAGCGCTACATCGCAAGTCTCGCCCAAGGACCGCGGACCGCGGCGTCTCTCTCTCT 461  
 Db 434 GGAGCGCTACATCGCAAGTCTCGCCCAAGGACCGCGGACCGCGGCGTCTCTCTCTCT 493  
 QY 462 ATTCGCGCTCTCGGCGCTCGAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521  
 Db 494 ATTCGCGCTCTCTCGGCGCTCGAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 551  
 QY 522 CCGGAGGTGAAGGCATGAGACACCGCGCGCGTCTCGCGCGCGTCTCTCTCTCTCTCT 581  
 Db 552 CCGGAGGTGAAGGCATGAGACACCGCGCGCGTCTCGCGCGCGTCTCTCTCTCTCTCT 611  
 QY 582 ATCCG-TTGCCGACGAGGGATGGAATCTGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 640  
 Db 612 ATCCGTTGCGGACGAGGGATGGAATCTGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 671  
 QY 641 GTGAGTCTGTGACAGTCTGCGGAGCTCAAGGCGTGTGAGGAGCAGACTCTGTCTCC 700  
 Db 672 GTGAGTCTGTGACAGTCTGCGGAGCTCAAGGCGTGTGAGGAGCAGACTCTGTCTCC 731  
 QY 701 ACTTCTACTCGCGTCAATCTCTGCGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTG 760  
 Db 732 ACTTCTACTCGCGTCAATCTCTGCGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTG 791  
 QY 761 ATGGGGTGGGTGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 820  
 Db 792 GTGGGGTGGGTGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 851  
 QY 821 TGATTGACCAACTTCTGCGCGGCTCTGCTGCTGATCAAGTTCGTTTCTGACATTG 880  
 Db 852 TGATTGACCAACTTCTGCGCGGCTCTGCTGCTGATCAAGTTCGTTTCTGACATTG 911  
 QY 881 TTCCGAGACTCTCATAGCGAGCTTTGGGTCCTTCTTAAGAGAGGATCTTGGA-TTT 939  
 Db 912 TTCCAGAGACTCTTAAGAGCGAGCTTTGGGTCCTTCTTAAGAGAGGATCTTGGA-TTT 971  
 QY 940 CGATTGTGATACAAAG 955  
 Db 972 TGATTGTATACAAAG 987

RESULT 4  
 AAD43512  
 ID AAD43512 standard; DNA; 923 BP.  
 XX AAD43512;  
 AC AAD43512;  
 DT 14-NOV-2002 (first entry)  
 XX Maize inositol polyphosphate kinase (IPPK) DNA #2.  
 XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
 KW nutritional value; animal feed; transgenic; gene; ds.  
 XX Zea mays.  
 OS  
 XX

PH Key Location/Qualifiers  
 FT CDS 53..736  
 FT /\*tag= a  
 FT /product= "Maize IPPK protein #2"  
 XX WO200259324-A2.  
 PN PN  
 XX 01-AUG-2002.  
 XX 09-JAN-2002; 2002WO-US003120.  
 PF 12-JAN-2001; 2001US-0261465P.  
 PR (PION-) PIONEER HI-BRED INT INC.  
 PA Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
 PI WPI; 2002-636540/68.  
 XX P-PSDB; AAE26194.  
 DR New inositol polyphosphate kinase polynucleotides and polypeptides,  
 DR useful in modulating phytic acid biosynthesis by decreasing phytate or  
 DR increasing non-phytate phosphorous to improve the nutritional value of  
 PT animal feed.  
 PT Claim 1; Page 60-61; 86pp; English.  
 XX The invention relates to novel inositol polyphosphate kinase (IPPK)  
 CC polypeptides and polynucleotides. Sequences of the invention are useful  
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
 CC increasing non-phytate phosphorous to improve the nutritional value of  
 CC animal feed, or to reduce the environmental impact of animal waste.  
 CC Polynucleotides of the invention are to produce transgenic plants with an  
 CC altered phenotype. IPPK proteins are used to screen compounds that  
 CC modulate their activity and raising anti-idiotypic antibodies. The  
 CC present sequence is maize IPPK DNA  
 XX Sequence 923 BP; 138 A; 326 C; 286 G; 173 T; 0 U; 0 Other;  
 SQ Query Match 62.3%; Score 837.4; DB 6; Length 923;  
 Best Local Similarity 97.5%; Pred. No. 1.9e-165;  
 Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;  
 QY 42 TCCCATACCATGTCGCGACCTCCACCGCGGAGCACAAGTCGCGCGCGCTC 101  
 Db 43 TCCCATACCATGTCGCGACCTCCACCGCGGAGCACAAGTCGCGCGCTC 102  
 QY 102 CGCCAGCAAGTGGGCGCGCTCATCGACGGCTCGCGGCTCTTTTACAAGCGCGTCCAGGC 161  
 Db 103 CGCCAGCAAGCGGCGCGCTCATCGACGGCTCGCGGCTCTTTTACAAGCGCGTCCAGGC 162  
 QY 162 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGGCGTTCCTCCGCCACGCGCGCT 221  
 Db 163 CGGCGACCGTGGGAGCAGAGGTGCGCTTCTATGAGGCGTTCCTCCGCCACGCGCGCT 222  
 QY 222 CCGGCGCGCATCCGAGACACCTTCTCCCGGTTCCACGGGACGCGACTCTCTCCGCCAC 281  
 Db 223 CCGGCGCGCATCCGAGACACCTTCTCCCGGTTCCACGGGACGCGACTCTCTCCGCCAC 282  
 QY 282 CGAGCGCGAGCCCGGGAGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 341  
 Db 283 CGAGCGCGAGCCCGGGAGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 342  
 QY 342 TCAGCGCGCTCTGCTCGAGACATCAAGATCGGCGCATCACTGAGGCGCATCGAGTTCGCC 401  
 Db 343 TGAGCGCGCTCTGCTCGAGACATCAAGATCGGCGCATCACTGAGGCGCATCGAGTTCGCC 402  
 QY 402 GGAGCGCTTACTCGCAAGTCTCTCGCGGAGCGCGGAGCGCGTCTCTGCTCGG 461  
 Db 403 GGAGCGCTTACTCGCAAGTCTCTCGCGGAGCGCGGAGCGCGTCTCTGCTCGG 462  
 QY 462 ATTCCGCGTCTCTCCGCGCTCGAGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 521



Db 807 GCGGGCTCTGCTAGCTGATCAAGTTGTTCTGACATTTGTTCCAGAGACTCTCTCAGCG 866  
Qy 901 CAGCCTTTGGGTCTCTTAAGAGAGGATCTGCA-TTTGATTTGATAACAAAG 955  
Db 867 CAGCCTTTGGGTCTCTTAAGAGAGGATCTGCAATTTTGAATTGATAACAAAG 922

RESULT 6  
AAD43518  
ID AAD43518 standard; DNA; 899 BP.  
XX  
AC AAD43518;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #5.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; gene; ds.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 89..424  
FT /\*tag= a  
FT /product= "Maize IPPK protein #4"  
XX  
FN WO200259324-A2.  
XX  
XX 01-AUG-2002.  
XX  
XX 09-JAN-2002; 2002WO-US003120.  
XX  
XX PR 12-JAN-2001; 2001US-0261465P.  
XX  
XX (PION-) PIONEER HI-BRED INT INC.  
XX  
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;  
FI P-PSDB; AAE26200.  
DR WPI; 2002-636540/68.  
DR P-PSDB; AAE26200.  
XX  
PT New inositol polyphosphate kinase polynucleotides and polypeptides,  
PT useful in modulating phytic acid biosynthesis by decreasing phytate or  
PT increasing non-phytate phosphorous to improve the nutritional value of  
PT animal feed.  
XX  
PS Claim 1; Page 73-74; 86pp; English.  
XX  
CC The invention relates to novel inositol polyphosphate kinase (IPPK)  
CC polypeptides and polynucleotides. Sequences of the invention are useful  
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or  
CC increasing non-phytate phosphorous to improve the nutritional value of  
CC animal feed, or to reduce the environmental impact of animal waste.  
CC Polynucleotides of the invention are to produce transgenic plants with an  
CC altered phenotype. IPPK proteins are used to screen compounds that  
CC modulate their activity and raising anti-idiotypic antibodies. The  
CC present sequence is maize IPPK DNA  
XX  
SQ Sequence 899 BP; 140 A; 312 C; 276 G; 171 T; 0 U; 0 Other;

Query Match 51.0%; Score 686; DB 6; Length 899;  
Best Local Similarity 88.7%; Pred. No. 8.4e-134;  
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;  
Qy 2 CACGAGGTCACTCCGTCAACCCCTCGCGCCATAGTCCCTTCCCATACCATGTCGACCC 61  
Db 45 CACCATCCGACCCGTCACCCCTTGCTCCATAG-----TCCCATACCATGTCGACCC 98  
Qy 62 TCCACCCGCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGTCGGGCCGC 121  
Db 99 TCCACCCGCGGAGCACCAAGTCGCGGCTCACCGGCTCCGCCAGCAAGTCGGGCCAC 158

Qy 122 TCATCAGCGCTCGGCTCTTCTAACAAGCGCTCCAGCGCGGCGACCGTGGGAGCACG 181  
Db 159 TCATCAGCGCTCTGGCTCTTCTAACAAGCGCTCCAGCGCGGCGACCGTGGGAGCACG 218  
Qy 182 AGTCCGCTTCTATGAGGGCTTCTCCGCCCAACGCGCGCTCCCGGCCGATCCGAGACA 241  
Db 219 AGTCCGCTTCTATGAGGGCTTCTCCGCCCAACGCGCGCTCCCGGCCGATCCGAGACA 278  
Qy 242 CTTTCTTCCCGGTTTCCAGGCAACGCACTCTCTCCCAACCGAGGCGGAGCCCGGGAGC 301  
Db 279 CTTTCTTCCCGGTTTCCAGGCAACGCACTCTCTCCCAACCGAGGCGGAGCCCGGGAGC 338  
Qy 302 CGCATCTCACCTCTCTCGAGCACTCTCTCGGGGGTTTTCAGGCGCCCTCGTCCGAG 361  
Db 339 CGCATCTGACCTCTCTCGAGCACTCTCTCGGGGGTTTTCAGGCGCCCTCGTCCGAG 398  
Qy 362 ACATCAAGATCGGCGCATCACGTCGCCACCGAGTTTCGCCGAGCCCTACATGCCAAGT 421  
Db 399 ACATCAAGATCGGTCGATCACGT----- 422  
Qy 422 GCCTGCCAAGGACCGGCGCATCACGTCGCCACCGAGTTTCGCCGAGCCCTACATGCCAAGT 481  
Db 423 -----GACCATGAGCGATCTGCTCGGATTCACGCTCTCGGGGTCC 463  
Qy 482 GAGTCTCGGCGCGGCGCTGTCGCGGAGCGGCGGCGGCGGAGGTGAAGGCCATGG 541  
Db 464 GAGTCTCGGCGCGGCGGCGCTGTCGCGGAGCGGCGGCGGCGGAGGTGAAGGCCATGG 523  
Qy 542 ACACCGCGGCGTCCGCGGCTGCTCCGCGCTACGTTGTCATCCG-TTGCAGACGAGGGG 600  
Db 524 ACATTGTCGGGCTCGCGGCTGCTCGGGCTGCTCATGTCATCGCTTGC CGGAGGGG 583  
Qy 601 ATGGAATGTCGCTCGCGGCGGCTGTCAGGAGGAAAGGTGGAATTTGTTCACAGCTG 660  
Db 584 ATGGAATGTCGCTCGCGGCGGCTGTCAGGAGGAAAGGTGGAATTTGTTCACAGCTG 643  
Qy 661 CGGAGCTCAAGGGGTGTTCCGAGGAGCAGACTCTGTCACCTTCTACTCGGGGTGAT 720  
Db 644 CGGAGCTCAAGGGGTGTTCCGAGGAGGCGGAGATGGGGTGGGGTGACGGTG 703  
Qy 721 CTTCTGGGCTATGATGCTGTCAGTCCGAGGCGGAGATGGGGTGGGGTGACGGTG 780  
Db 704 CTTCTGGGCTATGATGCTGTCAGTCCGAGGCGGAGGTCGGGGTGGGGTAAACAGTG 763  
Qy 781 AAGTGGTGAATTTGCCCCATGTCGCGGAGGTCGAGGTCGATGACCACTTCTCTG 840  
Db 764 AAGTGGTGAATTTGCCCCATGTCGCGGAGGTCGAGGTCGATGACCACTTCTCTG 823  
Qy 841 GCGGGGCTCTGCTCGCTGATCAAGTTGTTCTGACATTTCTCCGAGACTCTCTCATACG 900  
Db 824 GCGGGGCTCTGCTGATGATCAAGTTGTTCTGACATTTCTCCAGAGACTCTCTTAGACG 883  
Qy 901 CAGC 904  
Db 884 CAGC 887

RESULT 7  
AAD43519  
ID AAD43519 standard; DNA; 643 BP.  
XX  
AC AAD43519;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Maize inositol polyphosphate kinase (IPPK) DNA #6.  
XX  
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;  
KW nutritional value; animal feed; transgenic; ds.  
XX  
OS Zea mays.  
XX



```
PN WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 74; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
XX Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;
XX
XX Query Match 35.0%; Score 470.4; DB 6; Length 643;
XX Best Local Similarity 95.9%; Pred. No. 9.5e-89;
XX Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;
XX
Qy 42 TCCCATACCATGTCGACATCCACCGCGGAGCACCACCAAGTCGCGCACCGGCTC 101
Db 107 TCCCATACCATGTCGACATCCACCGCGGAGCACCACCAAGTCGCGCACCGGCTC 166
Qy 102 CGCAGCAGAGTGGGCGCGCTCATCGAGCGTCCGGCGCTTCTTACAGCGCGTCCAGGC 161
Db 167 CGCAGCAGAGTGGGCGCGCTCATCGAGCGTCCGGCGCTTCTTACAGCGCGTCCAGGC 226
Qy 162 CGGCGACCGTGGGAGCAGAGTGCCTTCTATGAGCGCTTCTCGGCCACCGCGCGT 221
Db 227 CGGCGACCGTGGGAGCAGAGTGCCTTCTATGAGCGCTTCTCGGCCACCGCGCGT 286
Qy 222 CCGCGCGCATCCGAGACATCTTCTCCCGCGGTTCCACGCGACGAGCTCTCCCGAC 281
Db 287 CCGCGCGCATCCGAGACATCTTCTCCCGCGGTTCCACGCGACGAGCTCTCCCGAC 346
Qy 282 CGAGCGCAGCCCGGAGCGCGATCTCTCACTCGTCTCGACGACCTCTCTCGGGGTT 341
Db 347 CGAGCGCAGCCCGGAGCGCGATCCGACCTCTCTCGACGACCTCTCTCGGGGTT 406
Qy 342 TCAGCGCGCTCGTTCGAGACATCAAGATCGGCGCCATCACTGCGGCCACCGAGTTGCC 401
Db 407 TGAGCGCGCTCGTTCGAGACATCAAGATCGGCGCCATCACTGCGGCCACCGAGTTGCC 466
Qy 402 GGAGCGCTTACATCCGAGTGCCTCGCCAGGACCGCGGGACCGAGCGTTCTGCTCGG 461
Db 467 GGAGCGCTTACATCCGAGTGCCTCGCCAGGACCGCGGGACCGAGCGTTCTGCTCGG 526
Qy 462 ATTCCGCTCTCCGCGCTCGAGTCTCGGCGCCCGCGCGCGCTGTCGCGAGCGAGCG 521
Db 527 ATTCCGCTCT--TGCCTCGAGTCTCGGCGCCCGCGCGCGCTGTCGCGAGCGAGCG 584
Qy 522 -CCGCGAGTGAAGSCCATGGACA--CCGCGCGCGCTCGCGCGCGTCTCTCCGG 570
Db 585 CCGCGCGGTTGAAGTATGACACCCCGTTCGNGNCGCGCGNGTCTCTCGG 635
```

---

```
RESULT 8
AAD43520
ID AAD43520 standard; DNA; 519 BP.
XX
XX AAD43520;
AC
XX 14-NOV-2002 (first entry)
DT
XX
XX Maize inositol polyphosphate kinase (IPPK) DNA #7.
DE
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
XX nutritional value; animal feed; transgenic; de.
KW
XX
XX Zea mays.
OS
XX
XX WO200259324-A2.
PN
XX
XX 01-AUG-2002.
PD
XX
XX 09-JAN-2002; 2002WO-US003120.
PF
XX
XX 12-JAN-2001; 2001US-0261465P.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI WPI; 2002-636540/68.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 75; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
XX Sequence 519 BP; 144 A; 109 C; 126 G; 125 T; 0 U; 15 Other;
XX
XX Query Match 22.2%; Score 298.4; DB 6; Length 519;
XX Best Local Similarity 89.4%; Pred. No. 8.5e-53;
XX Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;
XX
Qy 627 GTACGGAGGAAAGTGGAGTCTTGTCACAGCTGCGAGCTCAAGCGTGGTTCGAGGA 686
Db 2 GTACGGAGGAAAGTGGAGTCTTGTCACAGCTGCGAGCTCAAGCGTGGTTCGAGGG 61
Qy 687 GCAGACTCTGTTCCACTTCTACTCGGCGTCAATCTTCTGGGCTATGATGCTGTCAGT 746
Db 62 GCAGACTCTGTTCCACTTCTACTCGGCGTCAATCTTCTGGGCTATGATGCTGTCAGT 121
Qy 747 CGCAGCAGCGGAGATGGGCTGGGTGACCGTCAAGCTGGTGGACTTGGCCCATGTGGC 806
Db 122 CGCAGCAGCGGAGATGGGCTGGGTGACCGTCAAGCTGGTGGACTTGGCCCATGTGGC 181
Qy 807 CGAGGGTGATGGGTGATTGACCAAACTTCTCGGGCGGGCTCTGCTGCTGATCAAGTT 866
Db 182 CGAGGGTGATGGGTGATTGACCAAACTTCTCGGGCGGGCTCTGCTGCTGATCAAGTT 241
Qy 867 CGTTTCTGACATTGTTCCGAGACTCTCATACCGAGCCTTGGGTCCTTCTTAAGAGAG 926
Db 242 TGTTCCTGACATTGTTCCAGAGACTCTCTCAGACGAGCTTGGGTCTCTTCTTAAGAAA 301
```

```
QY 927 GATCCTGGCA-TTTCGATTGATAACCAAGCCCTACAAAGTTTCTCTCGGAAGAGCG 985
Db 302 GATCCTGGCATTTTCGATTGATAACCAAGGAANCATTTTCAGCTGCACAAAAAANCA 361
QY 986 CCTCCGA 992
Db 362 CCACTGA 368

RESULT 9
AAD43521
ID AAD43521 standard; DNA; 353 BP.
XX
AC AAD43521;
XX
XX 14-NOV-2002 (first entry)
XX
XX Maize inositol polyphosphate kinase (IPPK) DNA #8.
XX
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; ds.
XX
XX Zea mays.
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 75; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
XX Sequence 353 BP; 81 A; 78 C; 91 G; 96 T; 0 U; 7 Other;
Query Match 14.3%; Score 191.6; DB 6; Length 353;
Best Local Similarity 90.2%; Pred. No. 1.7e-30;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;
QY 667 CTCAGCGGTGTCGAGGAGCAGACTCTGTTCCACTTCTACTCGGGGTGATTTCTCTG 726
Db 1 CTCAGGCGATGGTTGGAGGAGCAGACTCTGTTCCACTTCTACTCGGGGTGATTTCTCTG 60
QY 727 GGCTATGATGCTGTCAGTCGACGAGCGGAGATGGGGTGGGGTGACGGTGAAGCTG 786
Db 61 GGCTATGATGCTGTCAGTCGACGAGCGGAGGATGGGGTGGGGTGAACAGTGAAGCTG 120
QY 787 GTGGACTTTGCCCATGTGGCCGAGGGGTGATGGGT--GATTGACCAAACTTCTGGGGC 844
Db 121 GTGGACTTTGCCCATGTGGCCGAGGGGTGATGGGTGATTTGACCAAACTTCTGGGGC 180

QY 845 GGCTCTGCTCGCTGATCAAGTTC--GTTTCTGACATTGTTCCGGAGACTCCTCATAGC-C 901
Db 181 AGCTCTGCTAGCTGATCAAGTTCGTTTCTTGACATTGTTCCAGANACTCCTTAGAGCC 240
QY 902 AGCCTTTGGTCCCTTCTTAAGAGA 925
Db 241 AGCCTTTGGTCCCTTCTTAAGAAA 264

RESULT 10
AAD43516
ID AAD43516 standard; DNA; 1195 BP.
XX
AC AAD43516;
XX
XX 14-NOV-2002 (first entry)
XX
XX Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.
XX
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
KW nutritional value; animal feed; gene; ds.
XX
XX Eucalyptus grandis.
XX
XX Key Location/Qualifiers
XX CDS 116..11048
XX FT /*tag= a
XX FT /product= "E. grandis IPPK protein"
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX
XX P-PSDB; AAE26198.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 68-70; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is Eucalyptus grandis IPPK DNA
XX
XX Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;
Query Match 13.7%; Score 184.6; DB 6; Length 1195;
Best Local Similarity 55.6%; Pred. No. 6e-29;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;
QY 56 CCGACCTCCACCGCGGAGACCAAGTCGCGGCCACCGCCTCGCCAGCAAGCTGG 115
Db 114 CCATGCTCAAGGTCCCGGATCATCAAGTCGCGGTCAACGGGGAGACGGGGGAAGCTGG 173
QY 116 GCCCGCTCATCGACGCTCGGCTTCTTACAAGCGCTTCCAGAGCGCGACCGTGGGG 175
```

```
Db 174 GGCCACTGGTGGATGATTGGGGCGCTTCTATAGCCCTCTCCAGAGCGATCATCGCGGAG 233
Qy 176 AGCAGAGGTGGCTTCTATGAGCGGTCTCCGCCACGCGCGTCCCGCCCGATCC 235
Db 234 ACACGGAAGTGGCTTTTACAGTCAATTTCTATTCCTCAATACCGAGATCCAGGTACATTC 293
Qy 236 GAGACACCTTCTTCCCGCGTTCACCGCACGCGACTCTCTCCACCGAGCGCGAGCCCG 295
Db 294 G---CAAAATCTTCTCGGTTTACCGAACTAAGACTAT---TGAGCGGTCTCATGGAT 347
Qy 296 GGGAGCGCATCTCACTCGTCTCCAGACCTCTCTCGCGGGTTCAGCGCCCTGCG 355
Db 348 CGGTCTCTCAACCTCACTCTGTTCTGGAGGATCTCGTCTGGGTTCGACGCAACCACTCTC 407
Qy 356 TCGCAGATCAAGATCGCGCCATCAGTGGCCACCGAGTTTCGCGAGCCCTACATCG 415
Db 408 TCATGGACATCAAGCTGGATTCAGAAACATGGTATCCGGAGGCTCTGAGAGGTACATCC 467
Qy 416 CCAAGTGCCTCGCCAAAGACCGCGGACACGAGCGTTCCTCGGATTCGCGCTCTCCG 475
Db 468 ABAAGTCTTAGAAGAGATCGAATAGCACAGCGTTTCATTGGGTTTGGATTTCTG 527
Qy 476 GCGTCCAGTGTCTCG---CCCCGAGGCGCGGTGTGGCGACGAGCGCGCGAGGTGA 532
Db 528 GGCTAAGGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAGAGTTGTT 587
Qy 533 AGGCATGGACACCGCGCGGTTCGCGCGTGTCTCCGCGGTAGCTGTCTATCC----- 585
Db 588 ATAGCTTTAATGCGGACGGTTCAGAGTCTCGGCTCTGAGGAAAGTTTGTTCCTCAACTGT 647
Qy 586 ----GTTGCGAGCGAGGATGACGTGGCGTCTCGCGCGGTGTACGAGGAAAG 640
Db 648 CTCTGGGTCCAAATGTGGATCCGATGTTTGTATGATCAAAAGTTTACTGTCAACCGG 707
Qy 641 GTGAGTCTTTGTCAAGCTGCGAGCTCAAGCGGTGGTTCGAGGACGACACTCTGTTC 700
Db 708 GTGGAATTTTGGCACAATTTGCTCAGCTGAAGGAATGGTTGAGGTTTCAGACGAATATC 767
Qy 701 ACTTCTACTGGGTCAATTTCTTGGGCTATGATGCTGTGACGTGCGAGCGCGGAG 760
Db 768 ACTTCTATTTCTGTCTACTCATTTATCTATATGACGAGGAGTCTGTCTTGACGCG--- 823
Qy 761 ATGGGGTGGGTGACGCTGAAGCTGTGGACTTTTGCCTATGTGCGCGAGGGTCAATGGG 820
Db 824 --TGTGCACACCCGAAAGTTAACTGTGTGACTTTGACATGTGATGATGGCCACGCG 881
Qy 821 TGATTGACCAAACTTCTGCGCGGGTCTGCTCGCTGATCAAGTTGTTTCTGACATTG 880
Db 882 TGATCGATCAACAATTTCTTGGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941
Qy 881 TTCCGGA 887
Db 942 CTGATGA 948
```

## RESULT 11

AAD43515 standard; DNA; 1105 BP.

XX AAD43515;

XX AAD43515;

DT 14-NOV-2002 (first entry)

XX Soybean inositol polyphosphate kinase (IPPK) DNA.

DE Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;

XX nutritional value; animal feed; transgenic; gene; ds.

XX Glycine max.

XX Key Location/Qualifiers

XX 12..851

XX /\*tag= a

FT

/product= "Soybean IPPK protein"

W0200259324-A2.

01-AUG-2002.

09-JAN-2002; 2002WO-US003120.

12-JAN-2001; 2001US-0261465P.

(PION-) PIONEER HI-BRED INT INC.

Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

WPI; 2002-636540/68.

P-PSDB; AAE26197.

New inositol polyphosphate kinase polynucleotides and polypeptides, useful in modulating phytic acid biosynthesis by decreasing phytate or increasing non-phytate phosphorous to improve the nutritional value of animal feed.

Claim 1; Page 66-67; 86pp; English.

The invention relates to novel inositol polyphosphate kinase (IPPK) polypeptides and polynucleotides. Sequences of the invention are useful in modulating the phytic acid biosynthesis by decreasing phytate and/or increasing non-phytate phosphorous to improve the nutritional value of animal feed, or to reduce the environmental impact of animal waste. Polynucleotides of the invention are to produce transgenic plants with an altered phenotype. IPPK proteins are used to screen compounds that modulate their activity and raising anti-idiotypic antibodies. The present sequence is soybean IPPK DNA

SQ Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;

Query Match 13.3%; Score 179; DB 6; Length 1105;

Best Local Similarity 55.6%; Pred. No. 8.8e-28;

Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;

Qy 61 CTCACCCCGCGAGACCAAGTCCGCGGCACCGCCCTCCGCGAGAGCTGGGCGG 120

Db 15 CTCGAAGATCCCGGAGCACCAGGTGGCGGCGCACAAAGCCAGGAACTCTGGGCCA 74

Qy 121 CTCATGACGGCTCCGCGCTCTTCAAGCCGCTCCAGCGCGCGCGAGCTGGGGAGCAC 180

Db 75 CTCGTGACGATTTGGAAAAATTTCAAGCCCTCCAGACCAACAAAGACGACGACACC 134

Qy 181 GAGTGGCTCTTATGAGGCGTT-----CTCGGCCCCACGCGCGCTCCGCGCGCGGCTATCC 234

Db 135 CGGGCTCCACCGNACTCTCTTTTACACCTCTCTCGCGCGCGCGCGGCTACTTCC 194

Qy 235 CGAGACACTTCTTCCCGCGGTTCACGGACCGGACTCTCTCCACCGAGGCGGAGCCC 294

Db 195 ATCGGTCTCTTCTCCCGGCTTTACGGCACCGCGCTCTCTGGACGCTCCGAGCGGTCC 254

Qy 295 GGGAGCGCGATCTCACTCGCTCTCTCGAGCCTCTCGGGGTTTCAGGCGCCCTGC 354

Db 255 GG---TCCCAACCTCTACCTGGTCTCTGGAGGACTCTCTCTGCGGCTACTCTCAAAACCTTCC 311

Qy 355 GTGCGACACATCAAGATCGCGCGCATCACTGGGCCACCGAGTTTCGCGGAGCCCTACATC 414

Db 312 GTATGACGTAAGATCGGCTCCAGAACCTTGGACCTGGGAGACTCCGAGGACTACATC 371

Qy 415 GCCAAGTCTTCGCAAGGACCGCGGACCAAGCGGCTTCTGCTCGATTCCGCGTCTCC 474

Db 372 TGCAAGTCTTGAAGAGGACAGAGTCTCTAGCTTGGCTTGGGTTTCAGAAATCTCG 431

Qy 475 GGGTCCGAGTCTGTCGCGCGCGGCGCGCTGTGGCGGACGCGCGCGCGGAGGTGAAG 534

Db 432 GGAGTCAAGGACTCTATCTCTCTCTGGGAACTTACAGGAAATCTCTCAAGTGTCTATCC 491

Qy 535 GCCATGACACCGCGCGCGTCTCGCGGCTGTCTCGGCGCT---ACGTGTCTATCGGTTGCC 591

Db 492 GCCCATGGTGGTGCATCTGTTCTCAACAAGTTCGTTTCTCTAATAATATCAACCATGAT 551  
QY 592 GACGAGGGAGTACGACTGTGCGCTCGCCGCGCGGTGTACGGAGGAAAAAGGTGGAGTCTTG 651  
Db 552 GATCATCATCCGATTTGGCTTTTGGCAACGGAGGTCTAC-----GGCGCGGTTTTG 602  
QY 652 TCACAGCTGGCGAGCTCAAGGCGTGGTTCGAGAGCAGACTGTGTTCCACTTCTACTCG 711  
Db 603 GAGCGGTTGCAGAAAGCTCAAGGACTGGTTGCGAGGTTTCAGACGGTGTATCACTTCTATTCT 662  
QY 712 GCCTCGATTCTTCTGGGCTATGATGCTGTGCAGTGCAGACGGCGAGATGGGGTGG 771  
Db 663 TGTTCTGTTCTTGTGTACGAGAGGATCTAGGAAA-----GGGAAAGTACCAAC 716  
QY 772 GTGACGGTGAAGTGTGAGCTTTGCCCCATGTCGCCGAGGGTGATGGGTGATGACAC 831  
Db 717 CCTCTGGTCAAACTCGTTGACTTTGCACACGTGGTGGACGAAACGGTGTCTATTGATCAC 776  
QY 832 AACTTCTGGCGGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACAT 878  
Db 777 AACTTCTGGGGTGGCTTTGTTCTTCTCATCAAGTTCCTCAAGGATAT 823

## RESULT 12

AAD43517  
ID AAD43517 standard; DNA; 1020 BP.

XX AAD43517;

XX 14-NOV-2002 (first entry)

DE P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.

XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;

KW nutritional value; animal feed; gene; ds.

XX Parthenium argentatum.

XX Key Location/Qualifiers

XX CDS 21..908

XX /\*Cag= a

XX /product= "P. argentatum IPPK protein #1"

XX WO200259324-A2.

XX 01-AUG-2002.

XX 09-JAN-2002; 2002WO-US003120.

XX 12-JAN-2001; 2001US-0261465P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX WPI; 2002-636540/68.

XX P-PSDB; AAE26199.

XX New inositol polyphosphate kinase polynucleotides and polypeptides,

XX useful in modulating phytic acid biosynthesis by decreasing phytate or

XX increasing non-phytate phosphorous to improve the nutritional value of

XX animal feed.

XX Claim 1; Page 71-72; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)

XX polypeptides and polynucleotides. Sequences of the invention are useful

XX in modulating the phytic acid biosynthesis by decreasing phytate and/or

XX increasing non-phytate phosphorous to improve the nutritional value of

XX animal feed, or to reduce the environmental impact of animal waste.

XX Polynucleotides of the invention are to produce transgenic plants with an

XX altered phenotype. IPPK proteins are used to screen compounds that

CC modulate their activity and raising anti-idiotypic antibodies. The

CC present sequence is P. argentatum IPPK DNA

XX SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;

Query Match 11.1%; Score 149.8; DB 6; Length 1020;

Best Local Similarity 53.2%; Pred. No. 1.1e-21;

Matches 456; Conservative 0; Mismatches 362; Indels 39; Gaps 5;

QY 61 CTCACCCCGCGAGCACCAAGTCGCGGCCACCGCGCTCCAGCGCGGACCGTGGGAGCAC 180

Db 24 CTCAGGCCCCAGATCATCAGGTGCTGACATGAAGCTGGGCTCGGAAGCTTGGCCCA 83

QY 121 CTCATCGAGCGTCCGGCTCTTCTACAGCGCTCCAGCGCGGACCGTGGGAGCAC 180

Db 84 CTCATGTGATGATTCAGCGCGGTTTACAAACACATGACGGGTGATTAACCGTGGTCA 143

QY 181 GAGGTGCGCTTCTATGAGCGGTTCTCCGCCACCGCGCTCCCGCCCGCATCCGAGAC 240

Db 144 GAAGTAGCCCTTTATGAATCATTTTCTTACATAATATTTCCAGAACACATACG--C 200

QY 241 ACCTTCTTCCCGGTTTCCACGGCAGCGACTCTCTCCCGCAGCGCGGAGCCCGGGAG 300

Db 201 AAATTTCTTCTATATATTTATGGCCACCAAAAT-----CATGAAGGCATCCACTGCTCT 254

QY 301 CCGCATCTCTCAGCTCGTCTCGACGACCTCTCGCGGGGTTTCAGCGCGCTCGCTCGCA 360

Db 255 GACCATCTCTCATGTTGTCAGATCTTAATCAGCTCATGTCAACCAATCTGTAATG 314

QY 361 GACATCAAGATCGGGCCATCACTGGGCCACCGAGTTCCCGGAGCCCTACATCGCAAG 420

Db 315 GACATCAAAATCGGTCCAGAACATGGGCGCCAGAGCTTCCGAGCGGTACATTCAGAA 374

QY 421 TGCCTCGCAAGAACCGCGGACCAAGCGTTCCTGCTCGGATTCGCGCTCTCCCGCGTC 480

Db 375 TGCCTAAAAAAGATAGGAAAGCACAAGTATTTCCATTGGGATTCAGGATCTCCGCGCTG 434

QY 481 CGAGTCTGCGCGCGCGGCGCGGCGGCTGTGGCGGAGG---AGCGCGCGGAGGTGAAGCC 537

Db 435 CAAGTCTATATCATGATGATGGGTCTTATAGCCCTCATAGAAATTCATGCGTAA 494

QY 538 ATGGCACACCGCGCGGCTCCCGCGCTGCTCCCGCGCTAGCTGTCTATCCGTTGCCGAGAG 597

Db 495 ACCGCGCCAGCTGATGTTAGACTACTTCTTAGGAAATTTGTTTCTTCTAACCGCTCTGCA 554

QY 598 GGGAT-----GGACTGTGCGCTCGCCCGCGGCGGTGTACGGA 633

Db 555 GAGATGGAAATGCGCACAGGCCCTAGGCCCGGATTTGTTCTTTAGCATCTTTTGTATGTT 614

QY 634 GGAAAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGTGGTTCGAGGAGCAGACT 693

Db 615 GGGCCTAATGGGATATTAGCTCAACTGATGGAAATTGACATGTTGAAGATCAACA 674

QY 694 CTGTTCCACTTCTACTCGCGCTCGATTTCTTCTGGGCTATGATGCTGCTGCGAGTCGAGCA 753

Db 675 ATTTACCACTTCTATGCTTGTGTTCTTTTGTTCATCTTTGAAAAGAGTTGGTGTAAAA 734

QY 754 GCGGAGATGGGGGTGGGTGACCGGTGAAGCTGCTGGGACTTTTCCCATGTGGCCGAGGTT 813

Db 735 GGT--GCTCGGTTCAAAACGACAGAAAGTCAAACTTATTGATTTTGTCTCATGTATCAGATGTT 791

QY 814 GATGGGTCATTTGACCAACACTTCTTGGGCGGGCTCTGCTCGCTGATCAAGTTCTGTTTCT 873

Db 792 AATGGTGTATTGATCACAATTTCTTTGGGTGGGCTCTGTTCTTTGATAAAGTTCAATTTCT 851

QY 874 GACATTGTTCCGAGAC 890

Db 852 GACATCTTTCGAGAC 868

RESULT 13

AAC39023

ID AAC39023 standard; DNA; 1168 BP.

XX AAC39023; PR 18-JUN-1999; 99US-0139462P.  
AC 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
DT 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
XX 23-JUN-1999; 99US-0140354P.  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS EPI033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128214P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147103P.  
PR 06-AUG-1999; 99US-0147166P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147835P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 17-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.

|   |              |   |
|---|--------------|---|
| PR  | 25-AUG-1999; | 99US-0150566P.  |
| PR  | 26-AUG-1999; | 99US-0150884P.  |
| PR  | 27-AUG-1999; | 99US-0151065P.  |
| PR  | 27-AUG-1999; | 99US-0151066P.  |
| PR  | 27-AUG-1999; | 99US-0151080P.  |
| PR  | 30-AUG-1999; | 99US-01511303P.   |
| PR  | 31-AUG-1999; | 99US-01511438P.   |
| PR  | 01-SEP-1999; | 99US-01511930P.   |
| PR  | 07-SEP-1999; | 99US-0152363P.  |
| PR  | 10-SEP-1999; | 99US-0153070P.  |
| PR  | 13-SEP-1999; | 99US-0153758P.  |
| PR  | 16-SEP-1999; | 99US-0154018P.  |
| PR  | 16-SEP-1999; | 99US-0154039P.  |
| PR  | 20-SEP-1999; | 99US-0154779P.  |
| PR  | 22-SEP-1999; | 99US-0155139P.  |
| PR  | 23-SEP-1999; | 99US-0155486P.  |
| PR  | 24-SEP-1999; | 99US-0155659P.  |
| PR  | 28-SEP-1999; | 99US-0156458P.  |
| PR  | 29-SEP-1999; | 99US-0156596P.  |
| PR  | 04-OCT-1999; | 99US-0157117P.  |
| PR  | 05-OCT-1999; | 99US-0157753P.  |
| PR  | 06-OCT-1999; | 99US-0157865P.  |
| PR  | 07-OCT-1999; | 99US-0158029P.  |
| PR  | 08-OCT-1999; | 99US-0158232P.  |
| PR  | 12-OCT-1999; | 99US-0158369P.  |
| PR  | 13-OCT-1999; | 99US-0159293P.  |
| PR  | 13-OCT-1999; | 99US-0159294P.  |
| PR  | 13-OCT-1999; | 99US-0159295P.  |
| PR  | 14-OCT-1999; | 99US-0159329P.  |
| PR  | 14-OCT-1999; | 99US-0159330P.  |
| PR  | 14-OCT-1999; | 99US-0159331P.  |
| PR  | 14-OCT-1999; | 99US-0159637P.  |
| PR  | 14-OCT-1999; | 99US-0159638P.  |
| PR  | 18-OCT-1999; | 99US-0159584P.  |
| PR  | 21-OCT-1999; | 99US-0160741P.  |
| PR  | 21-OCT-1999; | 99US-0160767P.  |
| PR  | 21-OCT-1999; | 99US-0160768P.  |
| PR  | 21-OCT-1999; | 99US-0160770P.  |
| PR  | 21-OCT-1999; | 99US-0160814P.  |
| PR  | 21-OCT-1999; | 99US-0160815P.  |
| PR  | 22-OCT-1999; | 99US-0160980P.  |
| PR  | 22-OCT-1999; | 99US-0160981P.  |
| PR  | 22-OCT-1999; | 99US-0160989P.  |
| PR  | 25-OCT-1999; | 99US-0161404P.  |
| PR  | 25-OCT-1999; | 99US-0161405P.  |
| PR  | 25-OCT-1999; | 99US-0161406P.  |
| PR  | 26-OCT-1999; | 99US-0161359P.  |
| PR  | 26-OCT-1999; | 99US-0161360P.  |
| PR  | 26-OCT-1999; | 99US-0161361P.  |
| PR  | 28-OCT-1999; | 99US-0161920P.  |
| PR  | 28-OCT-1999; | 99US-0161922P.  |
| PR  | 28-OCT-1999; | 99US-0161993P.  |
| PR  | 29-OCT-1999; | 99US-0162142P.  |
| Query Match 11.1%; Score 148.6; DB 3; Length 1168;              |              |   |
| Best Local Similarity 53.0%; Pred. No. 2e-21;                   |              |   |
| Matches 453; Conservative 0; Mismatches 369; Indels 33; Gaps 5; |              |   |
| QY  | 49           | ACCATTCGCACTCCACCGCGGAGACCAAGTCGCGGCGCACCGCGCTCCGCGCAGC 108     |
| Db  | 170          | AACAAGATGCAGCTCAAGATCCCTGAACATCAGGTTGCAGGACACATTGCTAAGACGGG 229 |
| QY  | 109          | AGCTGGCGCGCTCATCCAGCGCTCCGCCCTCTTACAGCGGCTCCAGCGCGCGAC 168      |
| Db  | 230          | AGCCTGGTCTCTCGTAGATGACAAGGTCGGTCTTCAAGCCACTTCAGGCGCATCTT 289    |
| QY  | 169          | CGTGGGAGCAGGAGTCGCTTCTATGAGCGCTTCTCCGCCCGCACCGCGCGTCCCGGCC 228  |
| Db  | 290          | CGTGGTGAATCGAGGTAAAGTTCTACGAATCTTTCTCTCAACACACAGAGGTTCCAGAA 349 |
| QY  | 229          | CGATCCGAGACACCTTTCTTCCCGGTTCCACGGACGCGATCTCTCCCGACCGAGGGC 288   |
| Db  | 350          | CACATCC---ATAGATATTTCGCGGTGTATCAGCGCACTCAA-----GCAGTT 394       |

|           |   |  |
|-----------|---|--|
| QY        | 289   | CAGCCCGGGAGCGCATCTCCTCGTCTCGAGGACCTCCTCGCGGGTTTCAGGCG 348          |
| Db        | 395   | GAAGTTCTGATGGAGCAGCCATGATGGTGTGGAAATCTTCTTGCAATACTCAAAA 454        |
| QY        | 349   | CCCTGCGTCGACATCAAGATCGGCGCCATCAGGTGCCACCGAGTTTCGCGGAGCC 408        |
| Db        | 455   | CCATCAGTAATGATGTTAAGATGGTTCACAGAACATGGTATCCTGATCATGAAGAA 514       |
| QY        | 409   | TACATCGCAAGTCCCTCCCAAGGACCGCGGACCAAGAGGTTCTGCTCGGATCCGC 468        |
| Db        | 515   | TACATCCAAAAATGTTTGAAGAAAGACACGGGTACCAACAACGGTGTCTATCGGGTTTCAGG 574 |
| QY        | 469   | GTCTCGGCGTCCGAGTCTGTCGCCGCCCGAGGGGCC---GTGTGGCGGACGAGCCCGC 525     |
| Db        | 575   | ATCTCTGGTTTCAAGTGTATGATCAAAAGATCGAGTTTCTGGAAGCCCGAGAGAAAG 634      |
| QY        | 526   | GAGGTGAAGGCCATGGACACACCGCGCGTCCCGCGTCTCCGCGCTACGTTGCA--- 582       |
| Db        | 635   | CTTCTTCGCGGGCTCGATGTAGATGGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 694   |
| QY        | 583   | -----TCGTTGCCGACGAGGGGATGGAATGTGCGCTCGCGCGCGGTGTACGGA 633          |
| Db        | 695   | AACTCACTTTTCGACACATGCGCTCGAAACCTGACTCTGCTTTTGCCTCGAGTGTTCACGCG 754 |
| QY        | 634   | GGAAAAAGGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGGCGTGGTTCGAGGAGCAGACT 693   |
| Db        | 755   | GTTTCCGACGGATCTTAACGCGATTGTCGGAACCTCAAGACTGTGTTGAGAACCAACG 814     |
| QY        | 694   | CTGTTCCACTTCTACTCGCGCTCGATCTTCTGGGCTATGATGCTGCTGCAGTC---GCA 750    |
| Db        | 815   | CTCTACCATTTCAACTCTTGTTCGATTTTAAATGGTCTATGAGAATGAATCCATCTTGAAG 874  |
| QY        | 751   | GCAGCGGAGATGGGGTGGGTGAGCGTGAAGCTGAGGACTTTCGCCCATGTGCGCCGAG 810     |
| Db        | 875   | GGAAATGATGATGATGATGCTAGACCACCAAGTCAAGCTGGTGGATTTTCTCATGTTCTGAT 934 |
| QY        | 811   | GGTATGGGCTGATTGACCACAACTTCTTGGCGGGCTCTGCTCGCTGATCAAGTTTCGTT 870    |
| Db        | 935   | GGTAATGGTGTCAATTGACCACTTCTTGGGTGGTCTTTGCTCTTTTCTTATAAATTCATT 994   |
| QY        | 871   | TCTGACATTGTTCCG 885  |
| Db        | 995   | CGTGAGATTCTTCAG 1009   |
| RESULT 14 |   |  |
| AAC33685  |   |  |
| ID        | AAC33685  | standard; DNA; 1104 BP.  |
| XX        | AC  | AAC33685;  |
| XX        | XX  |  |
| DT        | 17-OCT-2000   | (first entry)  |
| XX        | XX  |  |
| DE        | Arabidopsis thaliana  | DNA fragment SEQ ID NO: 3949.                                      |
| XX        | XX  |  |
| KW        | Hybridisation assay; genetic mapping; gene expression control;          |  |
| KW        | protein identification; signal transduction pathway; metabolic pathway; |  |
| KW        | promoter; termination sequence; ss.                                     |  |
| OS        | Arabidopsis thaliana.   |  |
| XX        | XX  |  |
| PN        | EP1033405-A2.   |  |
| XX        | XX  |  |
| PD        | 06-SEP-2000.  |  |
| XX        | XX  |  |
| PF        | 25-FEB-2000; 2000EP-00301439.   |  |
| XX        | XX  |  |
| PR        | 25-FEB-1999; 99US-0121825P.   |  |
| PR        | 05-MAR-1999; 99US-0123180P.   |  |
| PR        | 09-MAR-1999; 99US-0123548P.   |  |
| PR        | 23-MAR-1999; 99US-0125788P.   |  |
| PR        | 25-MAR-1999; 99US-0126264P.   |  |

PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0132456P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135829P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.



|   |   |
|---|---|
| PR 12-OCT-1999;   | 99US-0158369P.  |
| PR 13-OCT-1999;   | 99US-0159293P.  |
| PR 13-OCT-1999;   | 99US-0159294P.  |
| PR 13-OCT-1999;   | 99US-0159295P.  |
| PR 14-OCT-1999;   | 99US-0159329P.  |
| PR 14-OCT-1999;   | 99US-0159330P.  |
| PR 14-OCT-1999;   | 99US-0159331P.  |
| PR 14-OCT-1999;   | 99US-0159637P.  |
| PR 14-OCT-1999;   | 99US-0159638P.  |
| PR 18-OCT-1999;   | 99US-0159584P.  |
| PR 21-OCT-1999;   | 99US-0160741P.  |
| PR 21-OCT-1999;   | 99US-0160767P.  |
| PR 21-OCT-1999;   | 99US-0160768P.  |
| PR 21-OCT-1999;   | 99US-0160770P.  |
| PR 21-OCT-1999;   | 99US-0160814P.  |
| PR 22-OCT-1999;   | 99US-0160815P.  |
| PR 22-OCT-1999;   | 99US-0160980P.  |
| PR 22-OCT-1999;   | 99US-0160981P.  |
| PR 25-OCT-1999;   | 99US-0161404P.  |
| PR 25-OCT-1999;   | 99US-0161405P.  |
| PR 25-OCT-1999;   | 99US-0161406P.  |
| PR 26-OCT-1999;   | 99US-0161359P.  |
| PR 26-OCT-1999;   | 99US-0161360P.  |
| PR 26-OCT-1999;   | 99US-0161361P.  |
| PR 28-OCT-1999;   | 99US-0161920P.  |
| PR 28-OCT-1999;   | 99US-0161922P.  |
| PR 28-OCT-1999;   | 99US-0161993P.  |
| PR 29-OCT-1999;   | 99US-0162142P.  |
| Query Match 10.9%; Score 147; DB 3; Length 1104;<br>Best Local Similarity 52.9%; Pred. No. 4, 3e-21;<br>Matches 452; Conservative 0; Mismatches 370; Indels 33; Gaps 5; |   |
| QY 49   | ACCATGTCGAGACTCCACCGCGGAGCACCAAGTCGCGCGCACCGCGCTCCGCGCAGC 108       |
| DB 120  | AACAAGATGCGAGCTCAAGATCCCTGAACATCAGTTGTCAGGACACATTGCTAAAGACGGG 179   |
| QY 109  | AGCTGGGCGGCTCATCGAGCGCTCGGCGCTCTTCTAGAGCGGTCCTCAGCGGTCGAGCGCGAC 168 |
| DB 180  | AGCCCTGGTCTCTCGTAGATGACAAGGTCGGTCTTCAAGCCACTTCAGGCGGATTTCT 239      |
| QY 169  | CGTGGGAGCAGCAGGTCGCGCTTCTATGAGCGGTTCTCGGCCACCGCCCGCTCCGGCGC 228     |
| DB 240  | CGTGGTGAATCGAGGTAAGTTCTACGAACTCTTCTCTCAACACAGAGGTTCCAGAA 299        |
| QY 229  | CGCATCCGAGACACTTCTTTCGCCCGGTTCCACGGCACGCGACTCTCTCCCAACCGAGGCG 288   |
| DB 300  | CACATCC---ATAGATATTCCCGGTGTATCAGGCACTCAA-----GCAGTTT 344            |
| QY 289  | CAGCCCGGGAGCGGCATCTCACTCGTCTCGACGACTCTCTCGCGGGTTTCAGGCG 348         |
| DB 345  | GAAGGTTCTGATGGAGCAGCCATGATGGTGTGGAAAAATCTTCTTCGAGAATACTCAAAA 404    |
| QY 349  | CCCTCGGTCGACACATCAAGATCGCGGCATCACTGTCGCGCACCGAGTTTCGGCGGAGCCC 408   |
| DB 405  | CCATCAGTATGATGTTAAGATGGGTTTGAGNACATGGTATCTCTGATGATCTGAAGAA 464      |
| QY 409  | TACATCCCAAGTCCCTCGCCAAAGGACCGGGGACACGAGCGTCTCTCGGATTTCCGC 468       |
| DB 465  | TACATCCAAAAATGTTTGAAGAAGACACGGGTACCAACACCGTGTCTATCGGGTTTCAGG 524    |
| QY 469  | GTCTCGGGTCCGAGTCTGTCGCCCGCGAGGGCGCC---GTGTGGCGGACGGAGCGCCCG 525     |
| DB 525  | ATCTCTGGTTTTCGAAGTGTATGATCAAAAGATTCGAGTTTCTGGAAGCCCGGAGAGGA 584     |
| QY 526  | GAGGTGAAGCCCATGACACCGCGCGGTCCTCGCGGCTACGTTGTCA---582                |
| DB 585  | CTTCTTCGGGGTCCATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 644       |
| QY 583  | -----TCCGTTCCGACGAGGGATGGAATGTGCGCTCGCCCGCGCGGTGTAAGGA 633          |
| DB 645  | AACTCACTTTCGGACACTGGCTCGAAAACTGACTCTGCTTTTGCCTCGAGTGTTTACGGC 704    |

|                       |   |
|-----------------------|---|
| QY 634                | GGAAAAGGTGGAGTCTTGTCAAGCTCGCGAGCTCAAGGGCTGTTTCGAGGAGCAGACT 693          |
| DB 705                | GGTTCCACGGGATCTTAACGCGATTGCTGGAACTCAAGACCTGGTTCGAAACCAACG 764           |
| QY 694                | CTGTTCCACTTCTACTCGCGGTCGATTCTTCTGGGCTATGATGCTGCTGCAGTC---GCA 750        |
| DB 765                | CTCTACCATTTCAACTCTTGTTCGATTTTAATGGTCTATGAGAATGAATCCATCTTGAAG 824        |
| QY 751                | GCAGCGGAGATGGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGCCGAG 810          |
| DB 825                | GGAAATGATGATGATGCTAGACCACAAAGTCAAGCTGGTGGATTTGCTCATGTTCTTGAT 884        |
| QY 811                | GGTGATGGGTGATTGACCAACAACCTTCTGGCGGGCTCTGCTCGCTGATCAAGTTTGGTT 870        |
| DB 885                | GGTAATGGTGTCAATTGACCAATAAATCTTGGGTGGTCTTGTCTTTTATAAATTCATT 944          |
| QY 871                | TCTGACATTGTTCCG 885   |
| DB 945                | CGTGAGATCTTCAG 959  |
| RESULT 15<br>AAC38692 |   |
| ID                    | AAC38692 standard; DNA; 1243 BP.  |
| XX                    | AAC38692;   |
| AC                    | AAC38692;   |
| XX                    |   |
| DT                    | 17-OCT-2000 (first entry)   |
| XX                    |   |
| DE                    | Arabidopsis thaliana DNA fragment SEQ ID NO: 21894.                     |
| XX                    |   |
| KW                    | Hybridisation assay; genetic mapping; gene expression control;          |
| KW                    | protein identification; signal transduction pathway; metabolic pathway; |
| KW                    | promoter; termination sequence; ss.                                     |
| XX                    |   |
| OS                    | Arabidopsis thaliana.   |
| XX                    |   |
| PN                    | EP1033405-A2.   |
| XX                    |   |
| PD                    | 06-SEP-2000.  |
| XX                    |   |
| PF                    | 25-FEB-2000; 2000EP-00301439.   |
| XX                    |   |
| PR 25-FEB-1999;       | 99US-0121825P.  |
| PR 05-MAR-1999;       | 99US-0123180P.  |
| PR 09-MAR-1999;       | 99US-0123548P.  |
| PR 23-MAR-1999;       | 99US-0125788P.  |
| PR 25-MAR-1999;       | 99US-0126264P.  |
| PR 29-MAR-1999;       | 99US-0126785P.  |
| PR 01-APR-1999;       | 99US-0127462P.  |
| PR 06-APR-1999;       | 99US-0128234P.  |
| PR 08-APR-1999;       | 99US-0128714P.  |
| PR 16-APR-1999;       | 99US-0129845P.  |
| PR 19-APR-1999;       | 99US-0130077P.  |
| PR 21-APR-1999;       | 99US-0130449P.  |
| PR 23-APR-1999;       | 99US-0130510P.  |
| PR 23-APR-1999;       | 99US-0130891P.  |
| PR 28-APR-1999;       | 99US-0131449P.  |
| PR 30-APR-1999;       | 99US-0132048P.  |
| PR 30-APR-1999;       | 99US-0132407P.  |
| PR 04-MAY-1999;       | 99US-0132484P.  |
| PR 05-MAY-1999;       | 99US-0132485P.  |
| PR 06-MAY-1999;       | 99US-0132486P.  |
| PR 06-MAY-1999;       | 99US-0132487P.  |
| PR 07-MAY-1999;       | 99US-0132863P.  |
| PR 11-MAY-1999;       | 99US-0134256P.  |
| PR 14-MAY-1999;       | 99US-0134218P.  |
| PR 14-MAY-1999;       | 99US-0134219P.  |
| PR 14-MAY-1999;       | 99US-0134221P.  |
| PR 14-MAY-1999;       | 99US-0134370P.  |
| PR 18-MAY-1999;       | 99US-0134768P.  |
| PR 19-MAY-1999;       | 99US-0134941P.  |



|                 | Query Match           | 10.9%;  | Score 147;         | DB 3;      | Length 1243; |
|-----------------|-----------------------|---|--------------------|------------|--------------|
|                 | Best Local Similarity | 52.9%;  | Pred. No. 4.4e-21; |            |              |
|                 | Matches 452;          | Conservative 0;   | Mismatches 370;    | Indels 33; | Gaps 5;      |
| PR 26-OCT-1999; | 99US-0161361P.        |   |                    |            |              |
| PR 28-OCT-1999; | 99US-0161920P.        |   |                    |            |              |
| PR 28-OCT-1999; | 99US-0161992P.        |   |                    |            |              |
| PR 28-OCT-1999; | 99US-0161993P.        |   |                    |            |              |
| PR 29-OCT-1999; | 99US-01621142P.       |   |                    |            |              |
| Qy              | 49                    | ACCATGTCGAGCTCCACCGCGCGAGCACCAAGTCGCGGCCACCGCTCGCGCAGC 108            |                    |            |              |
| Db              | 257                   | AACAAGATGCAGCTCAAAAGTCCCTGACATCAGGTTCGAGACACATGCTTAAGACGGG 316        |                    |            |              |
| Qy              | 109                   | AAGCTGGGCGGCTCATCGACGGCTCGGGCTCTTTTACAAGCGCTCCAGCGCGCGAC 168          |                    |            |              |
| Db              | 317                   | AAGCTTGGTCTCTCGTAGATGACAAGGGTCGGTCTTTCGAAGCCACTTCAGGGCGATTCT 376      |                    |            |              |
| Qy              | 169                   | CGTGGGAGCAGAGGTGCGCTTCTATGAGCGTTCTCCGCCACGCGCGCGCTCCCGGCC 228         |                    |            |              |
| Db              | 377                   | CGTGGTGAATCGAGGTAAGTTCTACGATCTTTCTCTCTCAACACAGAGGTTCCAGAA 436         |                    |            |              |
| Qy              | 229                   | CGCATCGGAGACACTTCTTCTCCCGGTTCCACGGCACGGGACTCTCTCCGCCACGAGCG 288       |                    |            |              |
| Db              | 437                   | CACATCC--ATAGATATTCCCGGTGATCAGCGCACTCAA-----GCAGTT 481                |                    |            |              |
| Qy              | 289                   | CAGCCCGGAGCCGCAATCTCACTCGTCTCGACGACCTCTCTCGCGGGGTTTCAGCG 348          |                    |            |              |
| Db              | 482                   | GAAGGTTCTGATGGAGCAGCCATGATGGTGTGTGGAATAATCTCTTGAGAAATACTCAAAA 541     |                    |            |              |
| Qy              | 349                   | CCCTGCGTCGACAGACATCAAGATCGGCGCCATACAGTGGCCACCGAGTTCGCGGAGCCC 408      |                    |            |              |
| Db              | 542                   | CCATCAGTATGATGTTAAGATGGGTTTCGAAACATGGTATCTGATGATCTGAAGAA 601          |                    |            |              |
| Qy              | 409                   | TACATGCCCAAGTSCCTCGCCAAGGACCGGGGACACGAGCGTTCTGCTCGGATTCGCG 468        |                    |            |              |
| Db              | 602                   | TACATCCAAAAATGTTTGAAGAAAGACACAGGGTACCAACAACCGGTGTCATCGGGTTTCAGG 661   |                    |            |              |
| Qy              | 469                   | GTCTCCGGCGTCGAGTCTGTCGGCCCGAGGGCGCC---GTGTGGCGGACGAGCGCCG 525         |                    |            |              |
| Db              | 662                   | ATCTCTGGTTTCGAAGTGTATGATCAAAAAGAAATCGAGTTTCTTGAAAGCCGAGAGGAAG 721     |                    |            |              |
| Qy              | 526                   | GAGGTGAAGGCCATGAGCACACCGCGCGGTTCGCGCGTGTCTCGCGCGCTACGTGTCA--- 582     |                    |            |              |
| Db              | 722                   | CTTCTTCGCGGCTCGATGTAGATGAGACGAGATTGACTCTGAGAAAGTTTGATCATCT 781        |                    |            |              |
| Qy              | 583                   | -----TCCGTTGCCGAGGAGGGATGGACTGTGCGCTCGCGCGCGGGGTGTACGGA 633           |                    |            |              |
| Db              | 782                   | AACTCAGTTTCGGACACTGGCTCGAAACCTGACTCTGCTTTTGCTCGAGTGTTCAGCG 841        |                    |            |              |
| Qy              | 634                   | GGAAAAGTGGAGTCTTGTTCACAGCTCGCGAGCTCAAGGCGTGGTTTCGAGGAGCAGACT 693      |                    |            |              |
| Db              | 842                   | GTTTCCACGGGACTTTAAAGCAGTTGCTGGAATCAAGACCTGGTTCGAGAACCAACG 901         |                    |            |              |
| Qy              | 694                   | CTGTTCCACTTCTACTCGCGGTGCAATCTTTCTGGGCTATGATGCTGCTCGAGTC--GCA 750      |                    |            |              |
| Db              | 902                   | CTCTACCATTTCAACTCTTGTTCGATTTAATGGTCTATGAGAATGAATCCATCTTGAAG 961       |                    |            |              |
| Qy              | 751                   | GCAGGCGGAGTGGGGTGGGTGACCGGTGAAGCTGGTGGACTTTGCCCATGTGGCGGAG 810        |                    |            |              |
| Db              | 962                   | GGAAATGATGATGATGCTAGACCAACAAGTCAAGCTGGTGGATTGTGCTCATGTTTGTGAT 1021    |                    |            |              |
| Qy              | 811                   | GGTGATGGGGTGATTGACCAACAATTCTCGGGGGGCTCTGTCTCGCTGATCAAGTTCGTT 870      |                    |            |              |
| Db              | 1022                  | GGAATCGTGTCAATTGACCATTAATCTTGGGTGGTCTTTGGCTTTTGTCTTCAATAACTTCATT 1081 |                    |            |              |
| Qy              | 871                   | TCTGACATTTGTTCCG 885  |                    |            |              |
| Db              | 1082                  | CGTGAGATTCTTCAG 1096  |                    |            |              |